LIBRARY
STATE NORMAL COLLEGE
Y[illegible]

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS

HAROLD RUGG, *Chairman.*

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — L. M. TERMAN
FRANK N. FREEMAN — V. A. C. HENMON — BEARDSLEY RUML — E. L. THORNDIKE

VOL. XVII. No. 8 — NOVEMBER, 1926 — $4.00 a year. 60c per copy.

CONTENTS

TAPPING-RATE OF CHILDREN WHO TEST ABOVE 135 IQ (STANFORD-BINET). *Leta S. Hollingworth and Jane E. Monahan* 505

A PICTURE-PREFERENCE CHARACTER TEST. *Herbert Woodrow* . . 519

ON THE INADEQUACY OF THE PARTIAL AND MULTIPLE CORRELATION TECHNIQUE. *Barbara Stoddard Burks* 532

THE RELATION OF IQ TO POSITION IN FAMILY. *Grace Arthur* . 541

A NOTE ON SCALING TESTS. *Godfrey H. Thomson* 551

THE TECHNICAL VOCABULARY OF PSYCHOLOGY. *Charles L. Harlan* 554

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES 558

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 565

Published Monthly Except June to August by

Warwick & York, Inc.,

York, Pa. — Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania under the Act of March 3, 1879.

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. Title page and index are bound in December number.

Manuscripts for publication, books or other materials for review, and news items should be addressed to Harold Rugg, Chairman of the Editorial Board, 400 W. 118th Street, New York City.

The price of the Journal is $4.00 a year, payable in advance. Foreign postage is 30 cents extra. Single issues are 60 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for nonreceipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York, Inc.,
Baltimore, Md.

TEACHERS' HEADQUARTERS

European
Plan

Entirely
Fire-Proof

Centrally
Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

"I hail Bagley's book as the most important offering of the year. But you who are afraid that I boil over too easily, you who want to follow a master reasoner careful of his route, step by step, read Bagley on Determinism even through its last appendix." Educational Review, March, 1926.

DETERMINISM IN EDUCATION
By William C. Bagley

IN THIS book, the author reports in full the critical and constructive work that he has done during the past four years in connection with the fundamental questions raised by the efforts to measure native intelligence. Among the papers making up the volume are: Democracy and the IQ; Further Fallacies of Determinism; Do Good Schools Pay?; Education and the Moral Hazard; The Post-War Period as a Test of Universal Education; The Army Intelligence Tests and the Pro-Nordic Propaganda; Education as a Creative and an Equalizing Agency. The data presented in the studies heretofore published by the author have been greatly amplified and brought up to date. The recent investigations which have so signally confirmed the position that Mr. Bagley took in 1922 are described and summarized. An appendix includes ratings of the several states on a number of measures—such as economic efficiency, the production of leaders, intelligence, morality and criminality—as well as revised ratings of State school systems.

$2.20 + 10¢ *postage*

WARWICK AND YORK, Inc.
PUBLISHERS BALTIMORE

SCHOOL AND SOCIETY

A weekly journal covering the field of education in relation to the problems of American democracy.

The objects of SCHOOL AND SOCIETY are the advancement of education as a science and the adjustment of our lower and higher schools to the needs of modern life.

It encourages discussion of all problems from every point of view; it pays special attention to foreign affairs; it publishes accounts of educational events of the past week and research on fundamental problems.

Subscription $5.00—Single Numbers 15 Cents

[*SCHOOL AND SOCIETY will be sent to new subscribers for thirteen weeks for one dollar, which may be sent at the risk of the publisher.*]

THE SCIENCE PRESS
Grand Central Terminal - - New York, N. Y.

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XVII | November, 1926 | Number 8

TAPPING-RATE OF CHILDREN WHO TEST ABOVE 135 IQ (STANFORD-BINET)*

LETA S. HOLLINGWORTH

Teachers College, Columbia University

AND

JANE E. MONAHAN

Public School 165, New York City

In administering education for the intellectually gifted, the question is raised as to whether a young child can keep pace with less gifted, but older, children, in tasks which call primarily upon motor capacity; and if not, whether such considerations should affect policies of promotion. It is fairly common in the experience of the present writers for teachers to report that young children of superior intelligence, promoted much beyond the age-grade norms, are "deficient" in motor ability. It is said that they can not write well, can not manage wraps, can not perform the work assigned in manual training, and can not participate in athletic games. Thus a teacher complained of a 6-year-old of IQ 187, graded with 8- and 9-year-olds, that he was "inferior to the average child of his age in penmanship," forgetting, no doubt, that the average child of 6 years has developed no penmanship to serve as a basis of comparison.

* This report is rendered as part of the work of the joint committee in charge of special classes for bright children, previously described elsewhere[5,8]. The members of this committee are Mr. Jacob Theobald and Miss Jane Monahan, of Public School 165, New York City; Miss Margaret V. Cobb, Dr. Grace A. Taylor, and Prof. Leta S. Hollingworth, of Teachers College, Columbia University. The work has been carried on with the advice of District Superintendent John E. Wade, and in cooperation with the Division of Educational Psychology, of the Institute of Educational Research, of Teachers College. The stenographic and clerical labor of this study was provided for by funds granted through the Institute, by the Carnegie Corporation, of New York.

It is the purpose of this paper to report upon the maximal speed of voluntary movement in bright children, as compared with unselected children, using tapping as a test.

The Tapping Test

Comparatively little has been done with tests of motor capacity in children since 1905. Previous to that date several investigators had made use of tapping with right hand and with left hand, as a test of neuro-muscular quality. Bagley,[1] Bolton,[4] Binet,[3] Smedley,[10] and others worked with tapping as a test of motor power. Some of these investigators concluded that the relation between motor speed and intellect is positive; others, that it is indifferent. None reported an inferiority in tapping of children who surpassed their fellows in school work, and who were judged as intellectually "good" or "bright."

The present effort goes beyond previous studies in that the children tested for movement are of a known degree of intellectual superiority, by mental *tests*.

We conducted our test of tapping in accordance with the time limits set by Smedley, in order that comparison might be made with Smedley's norms. These norms were established upon public school children in Chicago, about 25 years ago, when the pitfalls of unintentional selection were not so widely understood as now. Still they remain the nearest approach we have to criteria of children's ability to tap.

We used the apparatus involving brass plate and stylus, with electric counter, taking care to have the contacts very clean, and employing three dry cells. Objections have been made to the use of the electric counter. For instance, Whipple[11] says:

"The electric counter is not reliable. Even with 10 or 12 cells of battery, it will miss a quick tap, which the graphic method will record. It follows that all results based upon the use of the counter are to be looked upon with suspicion."

Dewey, Child and Ruml[6] discarded tapping as a test, since "it was impossible to adjust the electric counter to record accurately."

Although aware of these objections, we found it impracticable to use the kymographic method of recording, and so used the electric counter. Our procedure was as follows. The child was seated at a comfortable height, before the apparatus, and was taught how to hold the stylus at an angle of about 45 degrees to the plate, the forearm resting on the table. The verbal instructions were these:

This test is to see how fast a child can tap. You must take the tapper (indicating stylus), and tap on this plate just as fast as you possibly can, this way (illustrating). As you tap, the machine will count, but pay no attention to that. Fix your whole mind on tapping just as fast as you can, till I say "stop." It is not how *hard*, but how *fast*, that counts; so don't tap this way (illustrating pounding), for it only wastes time. I shall say "Ready, go," and just as I say "go," you begin. Now, we'll see if you understand exactly. (A preliminary trial of about 20 taps with the preferred hand was here given, with a minute's rest after it.) When I say "go" begin instantly to tap.

The examiner then proceeded to the test, holding a stop-watch, released on the word "go." As the watch showed 30 seconds, the electric circuit was broken, thus stopping the counter. The number of taps registered was then recorded. In the test, each child tapped first with the right hand and then with the left, regardless of preferred hand.

If there is any constant error in the records, it is in the direction of causing the speed of movement to appear less than it really is. The enumerator never records taps which are not made, but can err only in failure to count quick, light taps. The children tested by us are, therefore, at least as swift in movement as the record shows. Our control group, tested with the same machine, and in all other respects in the same manner, will serve to show how bright children compare with the unselected in the number of taps which they are *able to register*, under identical conditions; thus answering the question which interests us.

Description of the Subjects

The 50 bright children tested in the present instance, were pupils in the special opportunity classes, at Public School 165, who have already been described elsewhere,[5,8]. They were selected from the public school population of New York City on two conditions, exclusive of all others, within the age limits prescribed. These conditions were (1) that they tested above 135 IQ (Stanford-Binet), and (2) that their parents permitted them to enter the classes.

There were also 31 children found eligible to the group, and invited to join it, whose parents declined the invitation, or withdrew them before the tapping test. Our committee is aware of the subtleties of unintentional selection, which enter wherever fewer than 100 per cent of cases in a category are studied, but it is our judgment that none of the factors which determined failure to join or adhere to our group had any bearing upon selection for ability to tap. The reasons given for declining or withdrawing were that parents preferred private

schools; moved from New York; did not wish to submit their children to educational experimentation; or desired that their children be advanced more rapidly through the elementary school than would be possible according to our plans.

To follow all of these "missing" children to the remote places whither they have now scattered, in order to obtain records in tapping, is not feasible. However, it has proved feasible to obtain measurements of physical size on 14 of these children, in addition to the mental measurements of all. In mental and physical measurements and in age, these "missing" children constitute a random sample of all found eligible. The only group differences which appear between the "missing" and the "included" are differences of sex and race. Among the "included" children, 30 per cent are girls, while among the "missing," 58 per cent are girls. Among the "included," 8 per cent are non-Jewish, while among the "missing" 26 per cent are non-Jewish, These differences of selection on the basis of sex and race have in themselves no invalidating effect upon our present findings, for the "missing" children are as bright and as large as the "included."

Since in mental and physical measurements they form a random sample of all found eligible, it is inferred that the children here tested are not in any respect a selection of the bright, as regards ability to tap. Of the 50 children who were members of our classes when the tapping test was given (in March, 1924) 46 were Jewish. This preponderance is due to peculiarities of segregation among pupils in New York City. School children are here divided among private schools other than parochial, parochial schools, and public schools. The segregations take place in such a manner as to gather Catholic children into parochial schools (especially the very bright, who attend largely where tuition is paid), while many Protestant families of superior economic status send their children to private schools, which are supported by endowment or tuition.[7] Our committee did not have access to private or parochial schools in selecting children for our classes. We have an indication of the extent to which very bright non-Jewish children are drawn off from the public schools, in the fact that we succeeded in collecting records from 100 non-Jewish children, testing above 135 IQ, in four private schools of comparatively small enrollment, where mental tests are given.* The great preponderance of Jewish children in our group does not, therefore, necessarily mean that the great majority of bright children in New York City are Jewish.

* To be reported in future by Margaret V. Cobb.

Whether tapping-rate in children is conditioned by race we do not know. Bere[2] found that 100 Jewish 10-year-olds fell below 100 Italian 10-year-olds in tests done with the hands, though surpassing the latter in tests of ability to deal with ideas.

A Control Group

Since we do not know (*a*) exactly what precautions may have been taken by Smedley against selection as regards intelligence,[2] whether there are race differences in manual performance, and (*b*) whether our apparatus is exactly comparable in function to that used by Smedley (because of the electric counter), we decided that a control group of children, tested by ourselves, would be necessary for valid interpretation of our records. We, therefore, drew a control group of 50 children from Public School 165, in the following manner.

For every bright child, a child matched in *sex*, *race* and *age* was drawn from the regular grades of the school, attention being given to these three factors alone, in matching. As regards age, it was not in every case possible to match exactly to the month, but we were able to bring the means of the groups within two weeks of each other (see Table IV). As regards intelligence, these control children probably have a central tendency of slightly above 100 IQ. The school is located in an excellent section of the City, and group tests given throughout the grades yield a central tendency of 110 IQ for the pupils. The very dullest children in the school are not represented, since they are segregated in the ungraded class, upon which we did not draw; while the very brightest are excluded by having been already placed in our special group. It is evident, therefore, that the control pupils are not ideally unselected, but we shall refer to them as unselected, since they are as nearly so as school children in the regular classes, at the ages studied, can be.

The 50 children of the control group, thus chosen, were tested in tapping by the same examiner, in the same room, with the same apparatus, during the same month, and with the same procedure exactly, as were the children of the bright group. Three of the control children were left-handed in tapping, matching three in the bright group.

Comparison with Smedley's Norms

Because investigators working subsequently may wish to make uses now unforeseen of data collected from a group thus highly selected, we have presented in Table I the complete data for each member of our bright group.

TABLE I.—PART A. SHOWING RECORDS IN TAPPING OF CHILDREN WHO TEST ABOVE 135 IQ (STANFORD-BINET). CHILDREN WHO ARE NOT JEWISH ARE INDICATED BY ASTERISKS. GIRLS ARE MARKED *g*. THE IDENTIFICATION NUMBERS OF THE CHILDREN ARE THOSE USED IN OTHER REPORTS BY THIS COMMITTEE

Child	Age	IQ	Taps in 30 seconds		Deviation from Smedley's norm		Index of unidexterity
			R	L	R	L	
1	10– 9	190	152	125	− 9	− 7	.82
2	10– 4	175	167	118	+ 6	−14	.71
3*g*	10– 3	167	157	134	0	+ 5	.85
4	10– 2	156	135	114	−26	−18	.84
5	11– 2	156	160	165	− 9	+24	.97
6	10–11	150	178	161	+17	+29	.90
7	10– 7	156	169	152	+ 8	+20	.90
8	9– 8	171	154	121	+ 3	− 6	.78
9	10– 1	160	149	108	−12	−24	.72
*10	10– 6	171	149	127	−12	− 5	.85
11	10– 5	162	178	109	+17	−23	.61
12*g*	10– 5	145	140	132	−17	+ 3	.94
13	9– 5	172	174	128	+23	+ 1	.74
14	11– 0	141	153	125	−16	−16	.82
15*g*	10– 7	147	169	131	+12	+ 2	.77
16	11– 3	150	171	149	+ 2	+ 8	.87
17*g*	10– 8	154	195	165	+38	+36	.85
18*g*	10– 8	152	200	169	+43	+40	.85
19	10–11	137	144	129	−17	− 3	.90
20	9– 4	163	138	110	−13	−17	.80
21	10– 4	146	147	136	−14	+ 4	.93
22*g*	9– 6	160	153	118	+ 4	0	.77
23	9–11	154	143	103	−18	−24	.72
24	10– 8	158	175	153	+14	+21	.87
25	9– 2	148	155	132	+ 4	+ 5	.85
26*g*	10– 1	170	158	120	+ 1	− 9	.75
27*g*	10– 5	164	160	117	+ 3	−12	.73
*29	9– 8	135	155	127	+ 4	0	.82
30	8–10	166	143	105	− 4	−12	.73
31*g*	10– 9	145	176	149	+19	+20	.85
32	10– 0	153	144	121	−17	−11	.84
33*g*	9–10	151	132	103	−17	−15	.78
34	9– 9	149	177	160	+26	+33	.90
35	9– 8	141	184	145	+33	+18	.79
*36*g*	9– 9	162	162	128	+13	+10	.79
37*g*	10– 5	139	132	165	−25	+36	.80
38	9–11	138	180	135	+29	+ 6	.75
*39	10– 8	140	155	130	− 6	− 2	.84
40	9–10	138	173	144	+22	+17	.83
41	10– 0	135	130	116	−31	−16	.89
42	9– 6	146	165	152	+14	+25	.92
43	10– 4	139	154	114	− 7	−18	.74
44	9– 1	155	140	149	−11	+22	.94
45*g*	10– 6	144	172	117	+15	−12	.68
46	10– 0	142	138	114	−23	−18	.83
47	9– 9	141	162	139	+11	+12	.86
49	9–11	141	151	120	0	− 7	.79
50*g*	9–10	138	138	135	−11	+17	.98
51*g*	9– 9	143	156	106	+ 7	−12	.68
52	9–10	154	132	118	−19	− 9	.89

In making our comparisons, we first matched each gifted child against Smedley's norm, for age and sex. According to Smedley, 9 years means past the ninth but not yet at the tenth birthday, and so forth, for each year. We, therefore, matched our children on that basis. For instance, Child 1, 10 years 9 months old, is matched with Smedley's 10-year norm for boys, and so forth. Smedley's norms for children of the ages studied are as follows, for number of taps registered in 30 seconds:

Age	Number tested	Boys		Number tested	Girls	
		R	L		R	L
8	31	147	117	31	146	117
9	60	151	127	44	149	118
10	47	161	132	48	157	129
11	49	169	141	48	169	139

In comparative distributions to show typical manual performance by right hand and by left hand, anomalies appear because of left-handed children. It occurs to us that in establishing norms for the hand, it would be desirable to standardize for *major* hand, instead of, or as well as, for right and left hand; or to standardize for right-handed and for left-handed, separately. We have matched right hand in every case against Smedley's norm for right hand, and similarly for left hand, regardless of preferred hand in the child matched. As stated, three of our 50 bright children made a major record with the left hand, in tapping.

Table II presents the frequency and amount of deviation from Smedley's norms. From the comparison, it is apparent that the gifted children studied do not differ reliably from the pupils upon whom Smedley found his norms, in respect to taps registered. The central tendency of the deviations clusters close around zero.

TABLE II.—SHOWING DISTRIBUTION OF DEVIATIONS FROM SMEDLEY'S NORMS, IN THE CASE OF CHILDREN WHO TEST ABOVE 135 IQ (STANFORD-BINET)

Amount of deviation (taps)	Frequency of children above 135 IQ	
	R	L
+45–35	2	3
35–25	3	2
25–15	5	9
15– 5	9	4
5– 5	10	10
– 5–15	10	12
15–25	8	10
25–35	3	0
Total.....	50	50

COMPARISON WITH THE CONTROL GROUP

The comparison with children in our control group is more interesting than the comparison with Smedley's norms, because we deal now with conditions known to be constant, as regards sex, race, age, apparatus and procedure. Table III presents the comparative distributions of performance, for bright children and for controls. Figure 1 shows these phenomena graphically.

TABLE III.—SHOWING DISTRIBUTION OF TAPPING RATE, FOR CHILDREN ABOVE 135 IQ, AND FOR CONTROL GROUP, UNDER IDENTICAL CONDITIONS

Taps registered in 30 seconds	Frequency of children above 135 IQ		Frequency of unselected children	
	R	L	R	L
205–200	..	..	1	
200–195	1			
195–190	1			
190–185				
185–180	1			
180–175	5	..	2	
175–170	5	..	3	1
170–165	3	1		
165–160	3	4	3	
160–155	5	1	1	
155–150	9	3	6	2
150–145	3	3	7	1
145–140	4	2	6	1
140–135	5	2	4	2
135–130	4	6	5	4
130–125	1	6	1	8
125–120	..	4	3	5
120–115	..	8	3	2
115–110	..	3	..	9
110–105	..	4	2	7
105–100	..	3	1	
100– 95	..	..	1	2
95– 90	..	..	..	2
90– 85	..	..	1	2
85– 80	..	..	..	1
80– 75	..	..	..	1
Total.....	50	50	50	50

It is seen that the children testing above 135 IQ are reliably swifter than the unselected children, as a group; and that this is true for both hands. Table IV gives the comparison of means, for the two groups.

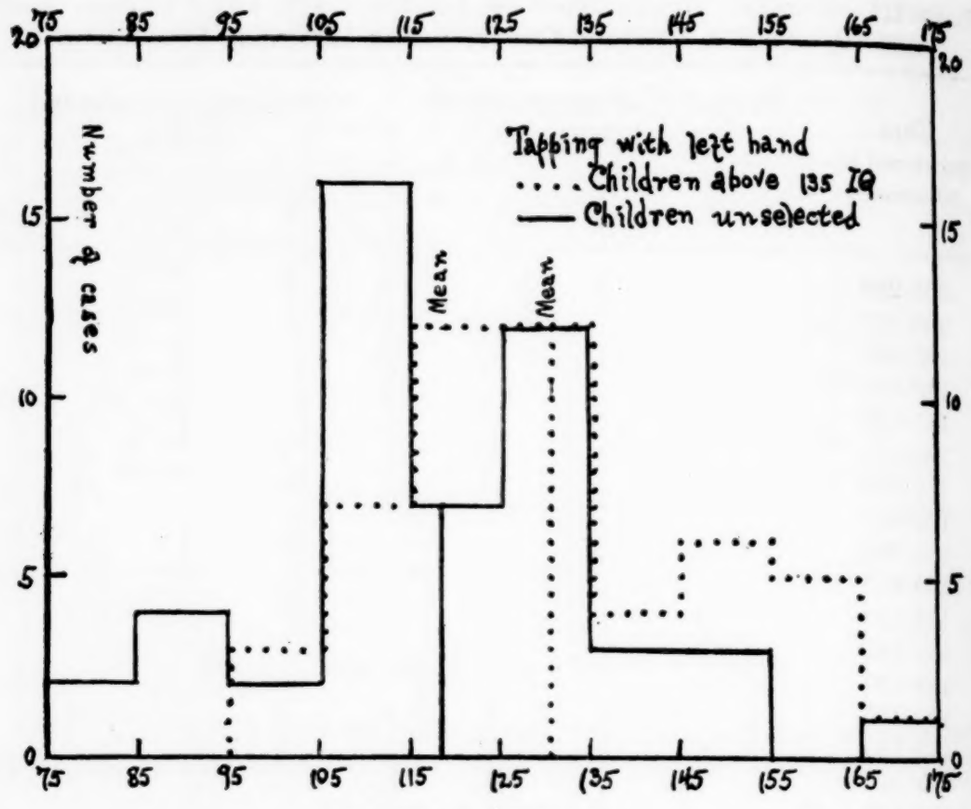

Tapping with left hand
.... Children above 135 IQ
— Children unselected
Mean
Mean
Number of cases
75 85 95 105 115 125 135 145 155 165 175
0 5 10 15 20

FIG. 1. PART A.

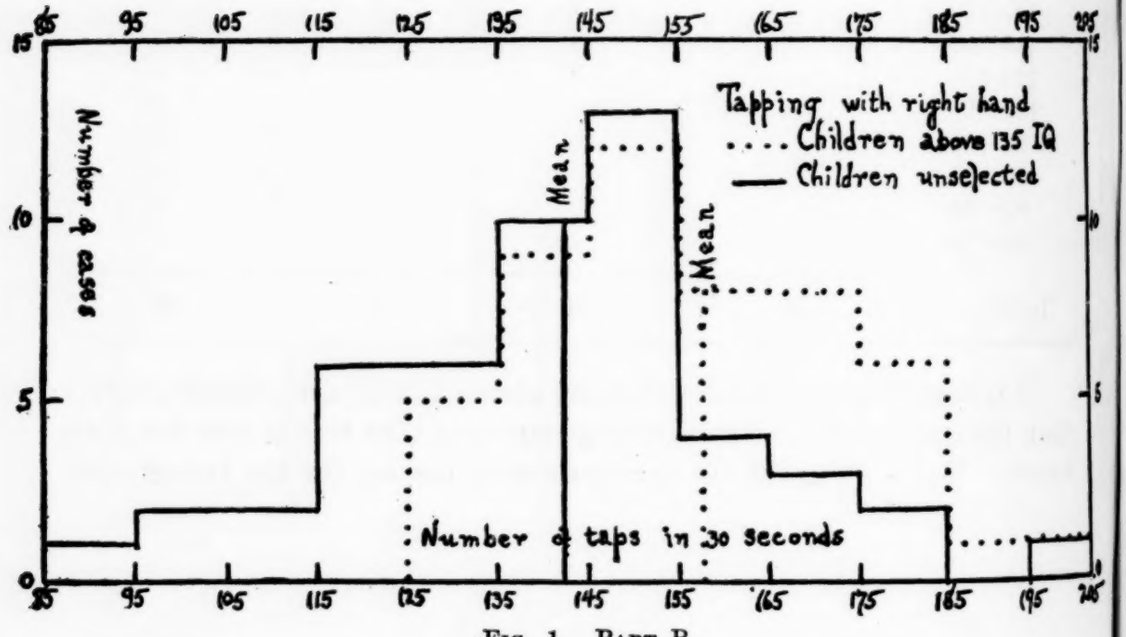

Tapping with right hand
.... Children above 135 IQ
— Children unselected
Mean
Mean
Number of cases
Number of taps in 30 seconds
85 95 105 115 125 135 145 155 165 175 185 195 205
0 5 10 15

FIG. 1. PART B.

FIG. 1.—Showing comparative distribution of taps registered by bright children, and by a control group unselected for intelligence.

TABLE IV.—SHOWING MEANS IN TAPPING RATE FOR CHILDREN ABOVE 135 IQ, AND FOR CONTROL GROUP UNDER IDENTICAL CONDITIONS; ALSO MEAN INDEX OF UNIDEXTERITY

Group	Average age in months			Average taps in 30 seconds						Average index of unidexterity		
				R			L					
	Av.	AD	PE_D	Av.	AD	PE_D	Av.	AD	PE_D	Av.	AD	PE_D
Above 135 IQ..	121.4	5.4	.645	157.4	12.8	1.53	130.9	15.0	1.79	.806	.066	.0079
Control.....	120.9	5.3	.634	142.7	16.6	1.98	118.2	14.4	1.72	.812	.070	.0084

NOTE.—It is of slight interest to know how much of the difference in the scores of the two groups may be attributed to the difference of two weeks in their mean ages. This may be inferred from knowledge of the amount which control children gain in speed, at these ages, in a year. In the right hand, control 10-year-olds surpass control 9-year-olds by 10 taps, and in the left hand by 18 taps. On this basis, it may be inferred roughly that the control group would have tapped about .4 of a tap more with the right hand, and about .7 of a tap more with the left, had they been two weeks older. About .5 of a tap in the difference between the scores of the two groups may thus be attributed to the discrepancy of two weeks in age.

INDEX OF UNIDEXTERITY

We mean here by *index of unidexterity* the ratio borne by minor to major hand $\left(\frac{\text{minor}}{\text{major}}\right)$. Table V gives the distribution of this index for bright group and for control group. Reference to Table IV will show the comparison of means, in this respect. There appears no reliable difference between the groups. The chances are even that the greater unidexterity shown here for the bright would disappear, with further sampling.

TABLE V.—SHOWING DISTRIBUTION OF CHILDREN ABOVE 135 IQ, COMPARED WITH A CONTROL GROUP, UNSELECTED FOR INTELLIGENCE, AS RESPECTS INDEX OF UNIDEXTERITY (RATIO OF MINOR TO MAJOR HAND)

Index	Frequency of children above 135 IQ	Frequency of control children
100–95	2	2
95–90	4	6
90–85	9	9
85–80	14	13
80–75	9	8
75–70	9	6
70–65	2	1
65–60	1	5
Total.....	50	50

Smedley found the *index of righthandedness* to be .82 at age 9, advancing to .89 at year 18. He says, "On the whole, there is a direct relationship between the degree of unidexterity and the intellectual progress of the pupil. At any given age of school life, bright or advanced pupils tend toward accentuated unidexterity, and dull or backward pupils toward ambidexterity." Mead[9] also found that stupid children tend toward ambidexterity, in grip.

CORRELATION BETWEEN SPEED AND STRENGTH AMONG THE BRIGHT

Measurements of ability to tap were correlated with measurements of grip, made in October, 1923, in our bright group. The Pearson coefficient of correlation for the right hand is .27; for the left hand, it is .05. Perhaps strength is more particularly a function of bone and muscle, while speed is more closely connected with quality of nerve.

CORRELATION BETWEEN SPEED AND HEIGHT AMONG THE BRIGHT

Measurements of standing height, made on the bright children in January, 1924, were correlated with the record of the major hand, in tapping. The Pearson coefficient is here −.10.

CONCLUSIONS

1. Children who test above 135 IQ (Stanford-Binet), with a mean at 151 IQ, and a range from 135 to 190 IQ, are distinctly superior, as a

group, to children unselected for IQ, as regards effective speed of movement in tapping. The bright register a greater number of effectual contacts than the unselected do, with identical opportunity.

2. The superiority of the bright is found in both hands.

3. No reliable difference is demonstrated, between the bright and the unselected, as regards the index of unidexterity. Such slight and perhaps chance difference as does appear is, however, in agreement with Smedley's statement that children of superior intellectual quality tend to more pronounced unidexterity.

4. Strength and speed of hand are slightly, if at all, correlated, among gifted children. Many of the strongest children in the group are relatively slow in movement, while the quickest are by no means always strong.

Our pupils of IQ above 135 were, on the average, 10 years 1 month and 1 week old, when the tapping test was given. They could then think with at least the cogency and persistence of average adults (median mental age being 15 years 1 month (Stanford-Binet); median score on Army Alpha being 108 points). They were about as tall and heavy as unselected children 11 years 6 months old. They could also use their hands in tapping and in gripping about as well as unselected children of 11 years 6 months. These statements apply to the children as a group. These bright children, selected originally by the test of intellect alone, have been found subsequently to be superior also in scholastic achievement,[5] in size, in strength,[8] and in speed. There is a great discrepancy, however, between amount of deviation in intellect and in scholastic achievement, on the one hand, and in size and motor capacity, on the other.

These facts obviously have direct implications for pedagogy. If graded with older pupils, whom they equal intellectually, children as young and as gifted as these will be well "out of their depth" as respects size, strength and speed. At 10 years of age, most of them can be advanced one year, and many of them two years, beyond the age-grade norms, without showing a "deficit" either in size or in motor capacity. This degree of advancement will not, however, take care of their intellectual needs, since they are, on the average, five years beyond the age-grade norms in capacity to assimilate ideas.

The alternatives of rational policy are, therefore, either to compromise between intellect and sensori-motor mechanisms in grading, or to segregate such children (where the population is dense) in special

classes, so that appropriate provision can be made for the exercise of all capacities, without the necessity of compromise.

Some of the findings of this investigation have long been apparent in an indefinite way to all who deal knowingly with young bright children. It is a fact of general observation that their physical and motor development by no means corresponds to their intellectual status, at a given age, in terms of birthday norms. The particular virtue of quantitative investigation is, however, two-fold. In the first place, it measures the amounts of the observed discrepancies, in terms of birthday norms, and of physical units. In the second place, it dispels the widespread illusion that bright children are "especially lacking" or "inferior" in manual capacity.

References

1. Bagley, W. C.: On the Correlation of Mental and Motor Ability in School Children. *American Journal of Psychology*, 1901.
2. Bere, M.: "Comparative Study of the Intelligence of Children of Recent Immigrants." Teachers College, Columbia University, 1924.
3. Binet, A. et Vaschide, N.: Epreuves de Vitesse chez les Jeunes Garçons. *Année Psychologique*, 1897.
4. Bolton, T. L.: The Relation of Motor Power to Intelligence. *American Journal of Psychology*, 1903.
5. Cobb, M. V. and Taylor, G. A.: Stanford Achievement Tests with a Group of Gifted Children. *Twenty-third Yearbook*, National Society for the Study of Education, Public School Publishing Co., Bloomington, Ill., 1924.
6. Dewey, E., Child, E. and Ruml, B.: "Methods and Results of Testing School Children." Dutton, New York, 1920.
7. Fenton, N. and Howard, L. S.: The Challenge of the Private School. *Journal of Educational Research*, 1924.
8. Hollingworth, L. S. and Taylor, G. A.: Size and Strength of Children Who Test above 135 IQ. *Twenty-third Yearbook*, National Society for the Study of Education, Public School Publishing Co., Bloomington, Ill., 1924.
9. Mead, C. D.: "The Relation of General Intelligence to Certain Mental and Physical Traits." Teachers College, Columbia University, 1916.
10. Smedley, F.: "Report of the Department of Child Study." Chicago Public Schools, 1900–01. (Original report not available. As presented in Report of the United States Commissioner of Education, 1920; and by Whipple.)
11. Whipple, G. M.: "Manual of Mental and Physical Tests." Warwick and York, Baltimore, 1914.

A PICTURE-PREFERENCE CHARACTER TEST

HERBERT WOODROW

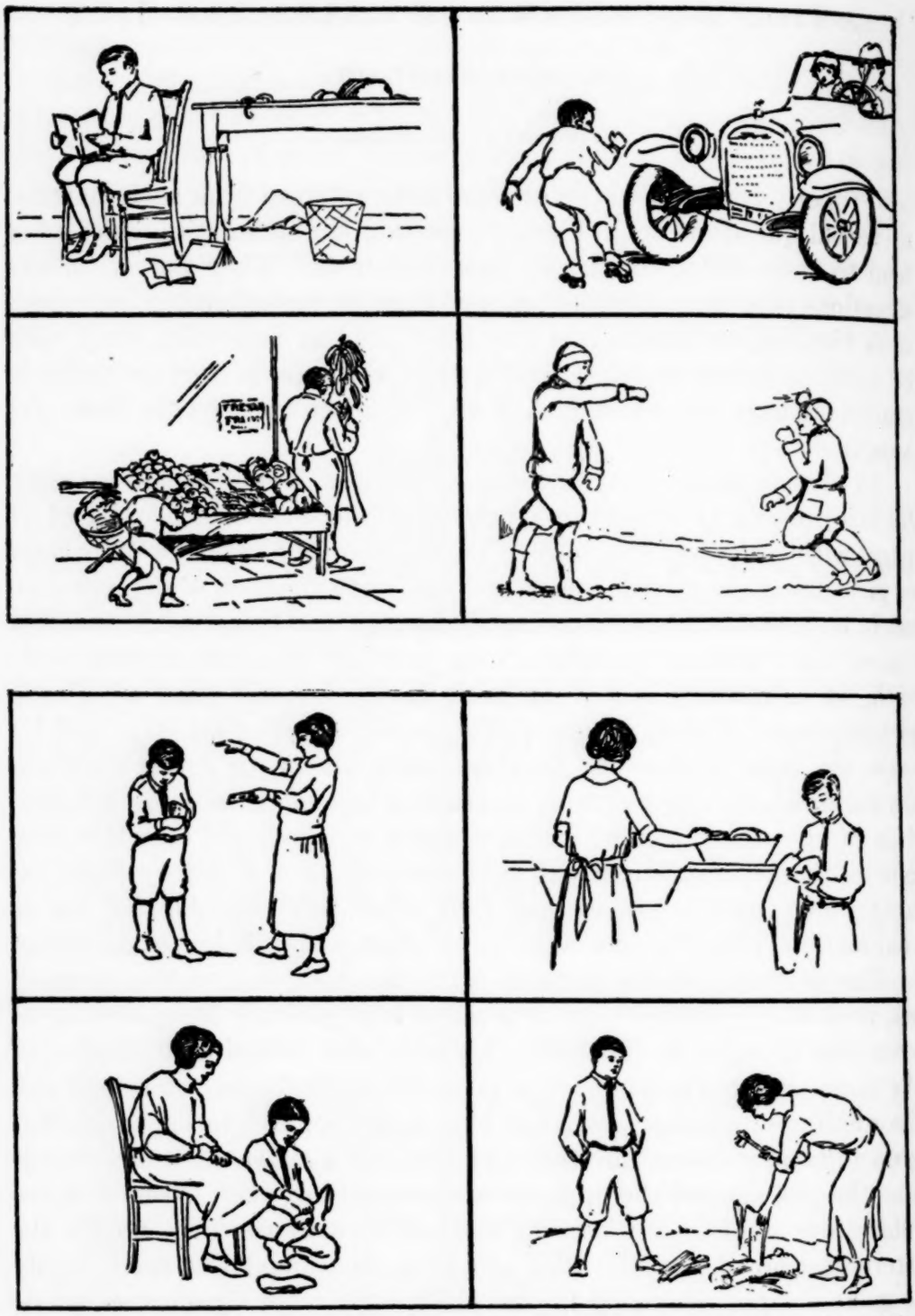
University of Minnesota

Properly constructed pictures are differentiated from such material as verbal questions or problems by certain characteristics which may lend to them value for use in character tests. They portray many situations in a more life-like manner than do verbal statements, and, since they do not necessitate the use of reading or writing, they may be used, as in the present investigation, with children in the primary grades, at ages for which few, if any, valid character tests have yet appeared.

In view of these considerations, a test was devised which consists, for the present, of 44 pictures, printed in the form of a booklet of 11 pages with 4 pictures on a page. Each picture shows one or more boys or girls doing something, and was drawn with the idea of depicting an act in a child which would ordinarily be regarded by an adult as either "good" or "bad." Samples of the pictures used are printed herewith, on a reduced scale. It need hardly be said that all of the pictures were of such a nature that no objection could be raised by even the most scrupulous to their being shown to young children. On the basis of these pictures, an attempt was made to devise for children of ages seven to nine inclusive a test which should correlate with general character sufficiently to be of value as one of a battery of several tests. The hope is entertained that other tests suitable for use in conjunction with the one here to be discussed will be forthcoming.

The grouping of the pictures, with the exception of the eleventh set, is of such a nature that each set of four pictures illustrates, or at least was thought to illustrate, desirable and undesirable degrees of the same trait. Thus one set is intended to illustrate orderliness and disorderliness; another good and bad ways for boys to fight; another good and bad manners; another a helpful, cooperative attitude towards a mother, as against the opposite of this attitude, etc. None of these rubrics are used in any way in connection with the test, nor in the interpretation of results. The grouping mentioned was used simply to make comparison easier for the child. No other significance whatsoever is attached to it.

The writer has used the test as a group test, and has introduced it to the children tested with the explanation that he was trying to find



Reproduction of two of the sets of pictures (reduced about two-thirds). Above, Set No. XII; below, Set No. VIII.

out what pictures little boys and girls like best. As a further introduction, the children were asked which of the pictures on the classroom walls they liked best, but their responses were not discussed. Then these instructions followed: "I am going to give you little picture books with four pictures on a page. I want you to look at each of the pictures on a page and see which one of the four you like the best. Mark it with a cross placed right on the picture (illustrated on blackboard). Then mark with two crosses the one you like second best, and with three crosses the one you like third best. The one you like best will not always be on the same part of the page, so you must look at all four pictures to be sure and see just what they show before you mark any of them." The instruction to look at all four pictures carefully before marking any of them was reinforced by having the children put down their pencils after turning to each new page and not letting them begin to mark until 30 seconds had elapsed. Also, each time they turned a page, they were reminded to look at each of the four pictures. A page was not turned until all the children had finished marking the last page, so that there was no time limit. The total time required to give the test varied from 12 to 15 minutes in Grade IV to about 20 minutes in Grade I. The test is thus short enough to allow of use in conjunction with other character tests.

It is easy to imagine instructions different from those employed, which aimed primarily to get the child to arrange the pictures in the order in which he "liked" them, without asking him to indicate the considerations determining his choice. Whatever ambiguity may lurk in these instructions, no child seemed unable, nor unduly reluctant, to carry them out. Care was taken to say nothing about "good" or "bad," the theory of the instructions being that if, because of any cuteness or other appealing characteristic in the picture, a child liked a "bad" picture, or else did not notice anything "bad" in it, he was a child whose attitude towards a similar situation in real life would be less desirable than that of one who discerned the "bad" feature and on that account was prejudiced against the picture in spite of whatever attractive features it may have possessed. Probably in actual life, each concrete instance of impoliteness, stealing, and so on, occurs under circumstances which lend to the act something attractive or which prevent the child from noticing, or at any rate particularly centering his attention upon, its "badness." In fact, one of the advantages accruing from the use of pictures may be that a child who, when asked directly whether he approves of stealing, would reply in

the negative, would, nevertheless, "like" a picture showing a boy cleverly appropriating an apple from the stand of a quaint dealer busy arranging his bananas. Thus, asking the child simply for an indication of liking stands a better chance of taking him "off his guard" than does direct questioning about the goodness or badness of the acts portrayed; and, unless this unsuspecting attitude on the part of the child be secured, there is little chance of arriving at any valid estimate of his true character.

Some mention should be made of the way the pictures finally used were selected, since, as it is important to observe, the success obtained depends upon the exact nature of the pictures. Nineteen sets of four each, drawn in ink, were first tried out on about 30 children selected by primary teachers as either the best or the worst in their class as regards general character. Each set of pictures had to fulfill certain demands. It had to contain pictures on which the children would disagree to a marked extent, and at least one picture which the "best" children would like decidedly more often than the "worst" children. It was almost impossible to predict before trial which sets would meet these demands. Some sets were discarded because all the children liked and disliked the same pictures. Others were discarded because each of the pictures was liked or disliked about equally often—that is, because the set was too hard. Nearly all sets were modified after the first preliminary trial, and were given a second preliminary trial on another group of children in a different school, before printing. A total of 120 pen-and-ink drawings was used in this preliminary study. Each of the sets finally chosen contains one or two pictures (those intended to illustrate good attitudes) which were definitely preferred to the other pictures of the set (those intended to illustrate bad attitudes) by the entire group of about 750 children, to whom the tests have now been given.

The pictures used were simply sketches, drawn partly from imagination, and partly from a series of over 50 photographs taken of children posed by the writer. The children used for posing possessed very little histrionic talent. Consequently, even apart from their sketchiness, none of the final drawings are actual reproductions from life. Simple sketches rather than completed drawings were used so as to avoid an over-abundance of irrelevant details.

Before the method of scoring was decided upon, the tests were given to a group of 648 children to determine in what order children themselves, as a group, liked the pictures. This group does not include

a group of nearly 100 used to determine the validity of the tests. Of course, the writer, or any other adult or group of adults, could have ranked the pictures in order of goodness; but it was preferred to use as a scoring norm the ranks assigned the pictures by children. The test was to be scored in terms of errors, and it was decided to regard as an error a deviation from norms established by the nine-year-old age-group—the highest of the age-groups for which the test is intended. The children tested for the purpose of establishing norms included all the children present at the time of testing in 22 school-rooms of the first, second, third, and lower fourth grades of three widely separated schools, two of them in Minneapolis, and one in St. Paul, and all three in supposedly representative neighborhoods. The number of children in each age group was as follows:

Age Group	*N*
6 years 0 months to 6 years 11 months inclusive	105
7 years 0 months to 7 years 11 months inclusive	170
8 years 0 months to 8 years 11 months inclusive	187
9 years 0 months to 9 years 11 months inclusive	142
10 years 0 months to 10 years 11 months inclusive	44

As already explained, the children ranked the pictures of each set of four in order from the one they liked best, to the one they liked least. From these rankings, a calculation was made of the average rank assigned each picture. This average rank of each picture is shown for the age-groups seven, eight, and nine in Table I. No picture has an average rank of one, for it could obtain this rank only if all the children liked it best. Likewise, no picture has an average rank of fourth or last.

The responses of each child were scored by giving him a certain demerit whenever he placed a picture in a position different from that corresponding to its average rank for nine-year-olds. For example, the average rank of the pictures of Set IV, letters being used as labels for the pictures, is as follows: *a*, 2.9; *b*, 2.5; *c*, 3.1; and *d*, 1.5. From these average ranks, it follows that the correct order for the pictures is *d*, *b*, *a*, *c*. If a child indicated a preference for them in any other order, he was given an error-score. If, for example, he placed picture *d* in the place which should be occupied by picture *a*, he was credited with an error equal to the difference in the average rank of *a* and *d* (in the case of Set IV an error equal to the difference between 2.9 and 1.5). In general, then, the error made by placing wrongly any

TABLE I.—AVERAGE RANK OF PICTURES ACCORDING TO CHILDREN'S PREFERENCES

N, for Age Group 7 = 170; for Age Group 8 = 187; for Age Group 9 = 142

Age-group.....	7:0 to 7:11				8:0 to 8:11				9:0 to 9:11			
Picture.....	*a*	*b*	*c*	*d*	*a*	*b*	*c*	*d*	*a*	*b*	*c*	*d*
Set I.....	2.3	2.4	2.7	2.5	2.5	2.2	2.7	2.5	2.8	2.2	2.7	2.3
Set II.....	1.9	2.9	3.0	2.2	1.7	2.9	3.1	2.2	1.9	3.2	2.9	2.0
Set III.....	2.9	2.5	2.5	2.1	3.4	2.7	2.2	1.7	3.5	2.9	2.2	1.5
Set IV.....	2.6	2.6	3.0	1.8	2.9	2.4	3.0	1.7	2.9	2.5	3.1	1.5
Set V.....	1.9	2.8	2.4	2.9	1.7	3.4	2.0	2.9	1.6	3.4	2.0	2.9
Set VI.....	2.8	2.1	3.0	2.1	3.2	1.9	3.0	1.9	3.4	1.8	3.0	1.8
Set VII.....	1.7	2.2	3.1	3.0	1.9	1.9	3.2	3.0	1.8	1.9	3.3	2.9
Set VIII.....	2.7	2.0	2.9	2.4	3.0	1.6	3.3	2.1	3.2	1.4	3.4	2.0
Set IX.....	2.4	2.0	2.5	3.1	2.6	2.2	2.1	3.2	2.7	2.2	1.9	3.2
Set X.....	2.5	1.9	3.0	2.5	2.9	1.7	3.3	2.2	3.0	1.6	3.4	2.0
Set XI.....	2.7	2.2	2.5	2.6	2.4	2.6	2.0	3.1	2.4	2.6	1.7	3.2

picture in a certain position was the difference between the average rank of the misplaced picture and that of the picture which would have been correct in that position. If the child interchanged two pictures differing widely in average rank, his error was large; while if he interchanged two which differed little, his error was small.

This method of scoring was chosen as the most practicable one for getting easily a reasonable weighting of the various errors made by any child. It is believed that this method solves some rather difficult problems, and that it might be employed to advantage in various researches where, instead of simply checking one of four responses, the individual arranges the responses in order of merit. However this may be, the method was chosen after a fairly thorough investigation of four different scoring methods, one of which involved no weighting of errors. This investigation was made by determining how picture preference demerit-scores agreed with various criteria of character to be mentioned below, and the results showed that the scoring method described above was superior to the others.

To facilitate the scoring, a tabular form was drawn up which showed all the possible orders in which the pictures could be ranked, and the corresponding demerit-scores. Thus, in Set IV, preferences for pictures in the following orders constitute the errors indicated.

SET IV

ORDER	ERROR
a b c d	3.2
a b d c	2.8
a c b d	4.0
a c d b	4.0 etc.

By means of this form the scoring can be done very quickly. If one has before him the order in which the child has ranked the pictures of each set, his demerit-score for the entire 11 sets may be determined in less than two minutes.

The possible scores on the entire test range by steps of 0.2 from 0.0 to 50.8. The obtained scores in one upper third grade ranged from 3.8 to 40.4. This range is sufficient to furnish a very satisfactory amount of differentiation between individual children.

In order to determine the reliability of the test, it was given a second time, three weeks after the first testing, to the children in the upper and lower second and the lower third grade of one of the schools. The Pearson coefficient of correlation obtained for the 95 children present at both times was 0.79. The average score on the first trial was 19.2, PE, 4.6; on the second trial, 17.3, PE, 5.5.

The correlation of the test with intelligence was calculated, by the Pearson formula, from data obtained on 214 children in the upper and lower divisions of the first three grades and the lower division of the fourth grade of the St. Paul school. These children had been tested with the Kuhlmann individual intelligence scale by a psychologist of the research department of the St. Paul public schools. Their mental ages were kindly furnished by the principal. The correlation between mental age on the Kuhlmann scale and score on the picture-preference test, with chronological age constant, for the 214 cases, was $-.24$. It should be remembered that the picture-preference scores are demerit-scores, so that the larger the score the poorer is the child's standing. The correlation $-.24$ means that the children of better intelligence made on the average the better picture-preference scores. Undoubtedly, intelligence is required for the observation and interpretation of the pictures. The correlation of the picture-scores with intelligence may be due, in part, to this fact. The complete data for this group of 214 children with respect to the correlation with intelligence are shown in Table II.

TABLE II.—DATA ON CORRELATION BETWEEN PICTURE-PREFERENCE SCORES AND MENTAL AGE

r, MA and picture-preference score, CA constant	−.24	PE .06
r, MA and CA	+.72	PE .03
r, MA and picture-preference score	−.48	PE .05
r, CA and picture-preference score	−.46	PE .05
Av. CA	8.3 yrs.	PE 0.9 yrs.
Av. MA	8.6 yrs.	PE 0.9 yrs.
Av. picture-preference score	19.7	PE 5.4
Range in CA	5.9 yrs.	to 12.7 yrs.
Range in MA	5.6 yrs.	to 12.5 yrs.
Range in picture-preference score	5.0	to 39.0

A correlation between picture-score and MA of −.24, when CA is constant, and of −.48, when both CA and MA vary over a range of seven years, shows definitely that the test is something of an intelligence test. This is further shown by the fact that the average demerit-score decreases with chronological age. The data on this point are presented in Table III.

TABLE III.—AGE NORMS FOR PICTURE-PREFERENCE SCORES

Age group	Av. score	PE	*N*
7:0 to 7:5	23.1	3.8	82
7:6 to 7:11	21.3	4.0	88
8:0 to 8:5	17.7	3.8	89
8:6 to 8:11	16.3	3.8	98
9:0 to 9:5	14.7	4.1	96
9:6 to 9:11	13.1	4.3	46

The preceding data immediately suggest the question whether a test which correlates with intelligence is not bound to be more a test of intelligence than one of character. The answer to this question depends upon what correlations the test shows with criteria of character, when intelligence is kept constant. Data on this matter follow, and they show that the correlations with character are by no means annihilated when calculated with mental age constant. Moreover, the existing evidence indicates that in children of the early ages here considered intelligence and character are by no means independent entities. The best criterion we could obtain of character, in a lower third grade, did itself show a correlation of +.48 with intelligence.

Terman observes that the correlation between IQ and average rating on 24 character traits, in first-grade children, was so high that "knowing only the average of these ratings one could roughly predict what the IQ would be."[1] The correlation was +.76. Ratings for a single trait, in a number of cases yielded correlations with IQ higher than +.50. For example, ratings for initiative and for persistence each showed correlations with IQ of +.53.

Even in older children, the correlation between character traits and intelligence is pronounced. Cady, who has published a battery of tests of "incorrigibility," states that in a population rated by teachers who were very sure of their judgment, the correlation between estimates of intelligence and estimates of corrigibility was +.66.[2] The correlation with group intelligence tests was +.32. May and Hartshorne report a correlation of +.40 between IQ and resistance of opportunities for cheating.[3] Raubenheimer calculated numerous coefficients of correlation between intelligence and tests of reliability, stability and healthy-mindedness, given to two age-groups, one of 11, and the other of 13 years. For the entire battery, he reports biserial *r*'s varying from −.05 to +.85. The individual tests showed correlations ranging from +.22 to +.77 with National Intelligence Test scores.[4] His overstatement test yielded a correlation with intelligence of +.64 (when corrected for attenuation, +.75). This correlation is particularly interesting, since Cady found the overstatement test which he employed to give a correlation of +.41 with his criterion. This is a higher correlation with the criterion than that shown by any other of the tests in his battery, and caused him to regard this test as probably his best.[5] On the whole, then, the evidence which investigators have so far gathered seems to indicate that an intelligence test is one of the best character tests of children yet devised, and that it is to be expected that a character test of young children will show a definite correlation with intelligence. If the existing data are valid,

---

[1] Terman L., M., after Dickson, V. E.: "The Intelligence of School Children." 1919, p. 61.

[2] Cady, V. M.: The Estimate of Juvenile Incorrigibility. *The Journal of Delinquency*, Monograph No. 2, 1923, p. 104.

[3] May, M. A. and Hartshorne, H.: First Steps toward a Scale for Measuring Attitudes. *Journal of Educational Psychology*, 1926, Vol. XVII, pp. 45–67.

[4] Raubenheimer, A. S.: An Experimental Study of Some Behavior Traits of the Potentially Delinquent Boy. *Psychological Monographs*, Vol. XXXIV, No. 159, 1923, p. 82.

[5] *Op. cit.*, p. 73.

it becomes highly improbable that there are any situations in life, where an individual's responses are determined exclusively by his traits of character and not affected by the degree of his intelligence.

The problem of determining the validity of any test of character is a hard one, because of the well-known difficulty of getting a reliable criterion with which to compare the test-scores. Teachers' rankings and ratings, as a rule, are open to serious question. Cady found it necessary to reject all but 150 cases out of some 700, rated by the teachers, because of the untrustworthiness of the ratings. The 150 cases retained consisted for the most part of boys who were rated as near the extremes of the distribution with respect to corrigibility. For these cases alone, did the teachers feel very sure of their judgments. Not only are any one teacher's rankings apt to be incorrect, but it is difficult, as a rule, to get independent rankings based upon prolonged observation from any teacher other than the one who has charge of the grade in which the children are found at the time of testing. While rankings were obtained, in the present study, from the teachers in one of the schools used, main reliance was placed on scores secured from one of the Minneapolis public schools in which an important and interesting experiment is now being conducted. In this school, the emphasis is placed on education in character; and to promote this emphasis, the only marks given are marks in character. Ten character traits, described by illustrations of their meaning in terms of concrete behavior, in a booklet furnished the teachers, are graded on a scale of five steps, designated by letters. The traits included are given the following names: reliability (including honesty), obedience, industry, social attitudes, self-control, judgment (*e.g.*, reasonable caution in crossing streets, judgment in hygienic matters, in disposal of time, etc.), punctuality, initiative, deportment and thrift. These traits, as defined by the illustrations in the booklet, constitute an excellent selection. In view of the fact that the teachers were thoroughly familiar with each of these rubrics, it was thought better to use the school ratings than to introduce any new rating scale. It seemed reasonable to believe, because of the fact that the marks in the character traits are the only marks the teachers give and because of the fact that the primary emphasis of the whole school is upon the observation and training of character, that teachers' ratings, in terms of the school's own rating system, would be definitely superior to ratings in an ordinary school, where the teachers are unaccustomed to this sort of marking. Unfortunately, no marks at all are given in this school until the third

grade. This necessitated a limitation of the study to the classes of the third grade. There were two upper third grades, designated $3A_1$ and $3A_2$ and one lower third grade, 3B. The fourth grade is beyond the age for which the test here concerned is intended. Only in the two 3A grades did marks exist for more than one semester.

Character marks were changed from letters to numbers, by substituting the numbers 1, 2, 3, 4 and 5 for the letters A, B, C, D and F in the case of the term marks, and the sum of the numerical scores of the 10 traits taken as a rating in general character. The score thus arrived at was used for calculating coefficients of correlation by the rank-order method. In the 3A grades, where two sets of term marks were at hand, the average of the two was used. Intelligence was measured by the Otis Group Primary Scale.

The coefficients of correlation, calculated by the rank-order method, are shown in Table IV. The correlation between rank in picture-preference score and rank in character rating, for the three grades, is as follows: $3A_1$, +.43; $3A_2$, +.41; 3B, +.38. The effect upon these

TABLE IV.—DATA CONCERNING RANK-ORDER COEFFICIENTS OF CORRELATION BETWEEN PICTURE-PREFERENCE SCORES AND CHARACTER RATINGS

School grade	$3A_1$	$3A_2$	3B
Picture-preference and character	.43 ± .11	.41 ± .11	.38 ± .11
Picture-preference and character (CA constant)	.47 ± .10	.49 ± .10	.38 ± .11
Picture-preference and character (MA constant)	.48 ± .10	.44 ± .10	.19 ± .12
Character and MA	.10 ± .13	.29 ± .12	.48 ± .10
Picture-preference and MA	.50 ± .10	.24 ± .12	.49 ± .10
Picture-preference and IQ	.46 ± .10	.38 ± .11	.42 ± .11
Picture-preference and CA	−.30 ± .12	−.37 ± .11	−.07 ± .13
Character marks (first term) and character marks (second term)	.81 ± .04	.88 ± .03	
Picture-preference and amalgamation of character and MA	.65 ± .08	.38 ± .11	.44 ± .10
Character and amalgamation of picture-preference and MA	.34 ± .11	.39 ± .11	.53 ± .09
Deviations from chronological age-norms in picture-preference and character	.36 ± .11	.39 ± .11	.36 ± .11
Picture-preference and character (Pearson *r*)	.38 ± .11	.46 ± .10	.25 ± .12
Average CA (years and months)	9:2	9:1	8:6
Average picture-preference score	13.7 ± 3.6	14.4 ± 4.8	17.8 ± 4.9
Number of children	28	30	29

coefficients of calculating them with mental age constant is not uniform. Those for the 3A grades are changed very slightly while that for the 3B grade sinks to +.19. The explanation of this difference between the grades lies mainly in the fact that in the case of the 3B grade the correlation between the character ratings, that is, the criterion itself, and mental age is as high as +.48, while in the two 3A grades it is much lower. The inverse correlation noted between CA and the picture-preference score is of course to be understood as holding only for these particular children as they came to be grouped by the school grading system. It would not hold for the entire child-population, as is clear from Table III, showing age-norms. It is due largely to this inverse correlation that the correlations between rank in picture-preference and that in character are raised, in the 3A grades, when calculated with CA constant. No corrections have been made for attenuation of the criterion, though a correction for this which would slightly raise the correlations could be legitimately made in the two 3A grades by utilizing the correlation between the first-term and second-term ratings.

Any test which shows norms varying with CA may of course be scored in two ways. One may use the score obtained irrespective of the CA of the child tested, as *e.g.*, a mental age score, or one may use a score obtained in relation to the age-norm, as in the case of the IQ. In calculating correlations for the three grades here studied, it appears somewhat questionable to use scores which take account of the child's CA. Just as one would use MA rather than IQ to determine the relationship of a test of intelligence to marks in school subjects, in a given school grade, so would one use raw scores in picture-preference rather than these scores taken in reference to CA norms, for comparison with teachers' marks in character. Nevertheless, for the sake of additional information, the picture-preference scores of each child have been tabulated as the deviations of the original scores from the norms of the CA age-group in which the CA of the child placed him. The correlations obtained by scoring the test in this way, in terms of the deviations from the age-norms, are given in the table. They are slightly lower than for the original scores, being +.36, +.39, and +.36 for the three grades, respectively.

In addition to the calculation of correlations, a comparison was made for the third grade as a whole between the average picture-preference scores of the best and worst thirds, as regards character ratings. The average character score, obtained as described above,

was approximately the same in each of the three classes, being 25.4, 23.5 and 24.0 for the $3A_1$, $3A_2$ and 3B grades, respectively. In each class the third with the highest and the third with the lowest character scores were chosen, and then the highest thirds put into one group and the lowest thirds into another. In this way two groups were secured of sufficient size to permit comparisons having some degree of reliability. The average picture-preference score for the third of the grade rating best in character is 11.7 ± 3.3 and for the third rating poorest in character 19.2 ± 4.6. The difference between the averages of the best and worst thirds is 7.5, which is 6.7 times the PE of the difference. As regards the amount of the difference, it may be described by the statement that 80 per cent of the worst third in character ratings made lower scores in picture-preference than the average of the best third, while 88 per cent of the best third scored higher than the average of the worst third.

Finally, it may be of interest to note the correlations obtained with six grades in the St. Paul school, where teachers were asked to rank the children as best they could in general character. The rank-order coefficients between teachers' character ranks and ranks in picture-preference scores were as follows:

Grade	ρ	N
1A	+.33	21
2B	+.09	26
2A	+.10	27
3B	+.45	31
3A	+.31	26
*3A	+.60	10

*3A pupils in a combined 3A and 4B section.

In conclusion, it should be stated that the preceding data are not regarded as definitely establishing much more than that the picture-preference technique shows promise. In spite of much time and care put upon their construction, the pictures now appear in need of a number of changes. After these changes have been made and after a further study of both different instructions and different groupings of pictures, it is intended more extensively to test the validity of this technique. The results so far obtained appear sufficient to justify the conviction that the test, if used in combination with one or two others showing not too high an intercorrelation, would have decided value.

ON THE INADEQUACY OF THE PARTIAL AND MULTIPLE CORRELATION TECHNIQUE

BARBARA STODDARD BURKS

Stanford University

Part I. In the Study of Causation

Logical considerations lead to the conclusion that the techniques of partial and multiple correlation are fraught with dangers that seriously restrict their applicability. In fact their attempted use in (*a*) isolating the causes which operate upon observed effects, and (*b*) defining the extent to which two measures involve common factors unique to themselves, often result in intepretations that are misleading and even untrue excepting in a few special types of situation. Only issues arising in the first field (*i.e.*, causation) will be discussed at this time. Consideration of the second field will be left for a subsequent paper (Part II).

Studies of Causation

Social and educational researchers probably find their most significant field in determining the causes of the effects which they observe. To know that physical defects do or do not cause inferior school achievement in children would be more important than to know merely that both tend to occur together; the knowledge that culture and wealth can or cannot raise children's IQ's would have implications much more fundamental to our social philosophy than any contained in mere association between intelligence and good environment; a knowledge of the exact contributions to the efficiency of teachers of the factors of training and native aptitude would be of immense value in outlining requirements for the teacher's certificate.

There is apparently a widespread impression that the methods of partial and multiple correlation are adequate substitutes for controlled experimentation in the solution of problems of causation. Not only are the techniques in wide favor for this purpose among research workers; but they are receiving definite sanction through editorials in the technical periodicals[2] and through their handling in current texts. One recent handbook,[8] for example, summarizes as a model in method a study[9] in which partial correlations have been conspicuously abused. Accordingly the significance to be attached to the issue is more than academic.

We may approach the problem inductively by first examining a hypothetical instance of the misapplication of partial correlation so extreme that it would scarcely be approached by any real situation. We wish to investigate, let us say, the effect of age of entering first grade upon subsequent rate of progress through the grades. We may take for our subjects an unselected group of 12 year old children. After noting for each child (*a*) the age at which he entered school; (*b*) his present grade; and (*c*) his rate of progress as measured by the average length of time it has taken him to complete each grade, we may find a substantial correlation between (*a*) and (*c*). Such a correlation would be reasonably expected if bright children tend both to enter school very young and to receive extra promotions.

Now let us render "present grade" constant by partialling it out. Then the correlation between age of entrance and time required to complete a grade necessarily falls to negative unity, since we are dealing now with children whose early entrance to first grade is exactly balanced by slow progress (*i.e.*, more than normal time spent in each grade); or whose late entrance to first grade is exactly balanced by rapid progress. The condition imposed that all children shall be in the same grade means that age of entrance must completely determine rate of progress. Yet it does not follow that the relationship between age of entrance and progress is absolute, irrespective of the "effect" of present grade upon progress. Indeed, present grade could have produced no *effect* whatever upon the rate of progress preceding it, since it is itself the direct *resultant* of age of entrance and rate of progress. The application of the partial correlation technique has thus been meaningless in so far as we are interested in getting at any true inter-dependence of age of entrance and rate of progress.

Next let us examine a less extreme situation. In an attempt to determine how much effect differences in environment have upon the development of children's intelligence we might collect data from an unselected group of families giving the intelligence test scores of a child and both his parents, and the family cultural status as measured on a specially devised scale. If the results checked with the trends indicated by various investigators, the correlations computed might not be far from the following:

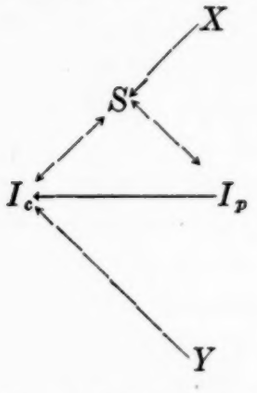
Mental age of mid-parent and intelligence of child	.45
Mental age of mid-parent and cultural status	.65
Intelligence of child and cultural status	.30

Using these hypothetical figures, let us calculate the partial correlation between intelligence scores of mid-parent and child. In current terminology we "hold constant" the cultural status, or "eliminate the effect" of cultural status.

Substituting in Yule's formula,[16]

$$r_{12.3} = \frac{r_{12} - r_{13}\, r_{23}}{\sqrt{(1 - r^2_{13})(1 - r^2_{23})}} = .35.$$

Should we then be justified in concluding that the real relationship between the intelligence of parents and their children after similarities induced by similar cultural surroundings had been discounted was measured by a correlation coefficient of only .35? Surely not, because we know that by the nature of partial correlation we have eliminated not only those portions of parents' and children's intelligence that may *result* from differences in cultural status, but those portions which *contribute* to cultural status as well. The situation is represented diagrammatically below. I_p represents intelligence of mid-parent; I_c, intelligence of child; S, cultural status; X, factors other than I_p or I_c contributing to S; and Y, combinations of genes not showing in the measurement of I_p contributing to I_c. It is readily seen that in apply-



ing the partial correlation formula to this situation we have "partialled out" too much, and that in any study of causation we are partialling out too much when we render constant factors which may *in part or in whole be caused by either of the two factors whose true relationship is to be measured, or by still other unmeasured remote causes which also affect either of the two isolated factors.*

The same generalization, of course, applies to studies of causation in which multiple correlation is used to find the "weights" to be attached to different factors in estimating a criterion; or in which multiple correlations are computed with different factors under

investigation successively dropped out to see how much each "contributes" toward estimating the criterion. While multiple correlations and weighting systems thus derived may be valid in themselves for purposes of prediction, serious misinterpretation often arises when cause is argued.

Perhaps the most notable misinterpretation in recent years of data based upon these methods is found in an argument put forward by Burt in 1921.[3] This argument, although questioned by Gates and La Salle with penetrating insight in 1923,[4] was quoted by Bagley[1] and by Thomson[11] in their latest books as an important link of evidence against the view that native ability is measured by intelligence tests. Holzinger and Freeman[5] have shown the argument to be entirely unwarranted.

Burt calculated a regression equation which predicted Binet-Simon scores of London school children by combining .54 schooling plus .33 "intelligence" (as measured on Burt's own reasoning test) plus .11 chronological age. He concluded that a child's mental age on the Binet-Simon tests "is a measure not only of the amount of intelligence with which he is congenitally endowed, not only of the plane of intelligence at which in the course of life and growth he has eventually arrived; it is also an index largely, if not mainly of the mass of scholastic information and skill which, in virtue of attendance more or less regular, by dint of instruction more or less effective, he has progressively accumulated in school."

Holzinger and Freeman, after demonstrating Burt's assumption of causal relationship to be untenable, capped their discussion by presenting a regression equation for predicting chronological age based upon Burt's own data. By this equation, chronological age of London school children could be estimated with minimal standard error by combining three factors in the following proportions: .15 Binet-Simon mental age, .51 schooling, and .03 mental age on the reasoning test. The authors submitted that to argue from this equation that chronological age was caused by schooling, Binet-Simon mental age, and reasoning power, would be quite as defensible as Burt's procedure in attributing causal significance to the factors from which he estimated Binet-Simon score. Their comments were climaxed by the ironical suggestion that perhaps the secret of eternal youth had been discovered at last!

The difficulty in the Burt study obviously lies in a confusion of the concept *prediction* with *cause*, and in overlooking the fact that in sys-

tems of correlated variables an effect yields as good an estimate of its cause as a cause yields of its effect.

Other interesting examples occur in the growing literature on prognostication of teaching success. Within the last four years Knight,[7] Somers[10] and Whitney,[12,13,14] have published studies upon that subject. Each of these investigators employed partial or multiple correlation in an attempt to find the degree to which each or all of certain measured traits or accomplishments contribute to teaching success. Knight, for example, sought among other things to establish the respective contributions of intellectual keenness, normal school standing, and ability to pass a progression test, to the teaching success of a group of elementary school teachers. To do this he obtained the second order partial correlation coefficient between each measure and teaching success with the other two measures constant. These partials were said by him to represent within the errors of sampling the "true relation" of the three factors to teaching success. But in considering, say, the partial correlation between intellectual keenness and teaching success we must remember that his procedure deletes not only the effect of normal school standing and grade in the professional test, but also that much of intellectual keenness which can be predicted from a knowledge of these two items. The correlated deviations in intellectual keenness and teaching success have been eliminated insofar as they are associated also with deviations in normal school standing and professional test scores. The real purpose of the treatment has thus been defeated.

Somers[10] and Whitney[14] examined the diminution of correlation between teaching success and factors such as intelligence, personality ratings, school marks and practice-teaching as partial correlation coefficients of higher and higher order were successively computed. They also noted the variations which the multiple correlation with the criterion underwent as each factor was in turn subtracted from the battery and then replaced. Whitney, it is true, pointed out in one of his papers[12] that his fifth order partials merely "represent in each case the correlation with teaching success of what is unique to each of the independent variables," but in the same paper he ranked the factors in order of their "relative importance" to teaching success according to their weights in a regression formula and their contributions to the magnitude of the multiple correlation of the battery with the criterion. In a paper appearing somewhat later[13] he suggested that the method of partial correlation as he has applied it would be useful in planning

the curriculum for teacher-training. The courses showing the highest partial correlations with later success, he reasoned, should receive most consideration in planning the curriculum.

A curriculum analysis based upon these methods would be subject to unavoidable pitfalls. Other things being equal, regardless of their real contributions to successful teaching, the courses appearing to make the slightest contributions to teaching success would be those in which the intercorrelations of students' marks were highest. It is further clear (although this point is aside from the discussion of the partial correlation), that the untenable assumption has been made that the correlation between marks in a subject on one hand and teaching success on the other is a valid index of the contribution of that subject to teaching merit. To illustrate, consider a subject, such as the multiplication table, so essential that no teacher could function without it, yet so universal that practically no teacher enters the field or even the normal school without a mastery of it. The correlation between knowledge of the multiplication table and rating on teaching success would be negligible, yet we should hesitate to send out teachers who were not equipped with this fundamental skill. Is it not probable that there are presented in normal school courses skills, theories, or cardinal principles so thoroughly mastered by all students as not to show up in correlations of marks with teaching success, but without which it would be most unfortunate to certify teachers? May there not also be courses which contribute relatively little to teaching merit, but which show high correlations with teaching success merely because they demand the same group of abilities demanded in actual teaching?

Returning to our main problem, and noting briefly a few additional examples from diverse fields, we may touch upon Reavis' study[9] of factors controlling attendance in rural schools. Reavis obtained data for over 5000 children upon the following: Attendance, distance from school, age-grade status, quality of work, merit of teacher, merit of school plant, and kind of community. The fifth order partial correlations of these factors with attendance are presented as representing their true contribution to the criterion. However, on the face of it, some of the fifth order partials must represent instances in which part of the "cause" actually under inspection has been eliminated due to its connection with remote causes which affect it in common with other factors that have been partialled out. Also, some of the factors are undoubtedly connected by direct paths of influence, so that some

of them might properly be partialled out, while others might not. Quality of work, for example, is obviously a contributing cause to age-grade status. Hence it would probably be a valid procedure to partial out quality of work when the contribution of age-grade status to attendance was being considered; but not to partial out age-grade status when the contribution of quality of work was being considered.

As a final representative example, we may cite a portion of Bagley's *Determinism in Education.*[1] In his Chapter IV, *Do Good Schools Pay?*, he attempted through the correlations of state rankings upon school indices with certain measures of prosperity, culture, and leadership, to show that good schools are responsible for the desirable attributes of modern life. Taking the partials of the intercorrelations between these factors with "school index" rendered constant, he found them much reduced in magnitude from the raw values of the intercorrelations. He interpreted these findings as meaning that school opportunities had exerted a causal influence upon the other factors. But in drawing conclusions from partials with school index rendered constant he has begged his question. It may or may not be true that the intercorrelations of state rankings in leadership, prosperity and culture are due to the relative excellence of their educational systems; but Bagley's data have not confirmed a causal relationship. The same reductions in the values of the intercorrelations of these measures could be obtained if: (1) School facilities were the direct outcome of them rather than their cause, or if (2) school facilities resulted from remote causes that determined also the other measures.

The question at last arises as to whether partial or multiple correlation can ever be fully defended as instruments in the study of causation. Surely, as Kelley has pointed out,[6] there is not more justification for inferring causation in partial correlation than in raw correlation. Nevertheless, there are situations in which a variable is indisputably a cause rather than an effect. In one obvious type of situation, chronological age is such a variable. The many studies which have employed partial correlation technique to eliminate the contribution of maturity to correlated measures are apparently on safe ground, provided they have published the precise age ranges of their subjects. However, the caution sounded by Yule[16] in his chapters on partial correlation and normal correlation should not be lost sight of: Namely, that excepting in the relatively rare situation of normal correlation, the partial coefficient cannot be taken as a measure of the correlation

between variable 1 and 2 for *every* constant value of variable 3, but only as a sort of average correlation.

Finally, a paper on *Correlation and Causation* published by Sewall Wright in 1921[15] should be cited. In this study, Wright criticizes the partial correlation technique and develops a substitute which he calls a method of path coefficients for "finding the degree to which variation of a given effect is determined by each particular cause." "The method," he writes, "depends on the combination of knowledge of the degrees of correlation among variables in a system with such knowledge as may be possessed of the causal relations. In cases in which the causal relations are uncertain, the method can be used to find the logical consequences of any particular hypothesis in regard to them."

Wright illustrates his technique chiefly by examples from agriculture and animal study, but his method is quite general. The writer recalls no study in educational or social statistics that has made use of it, but suggests that it may prove to have some applicability. Its availability is, of course, restricted by the rarity with which we have definite knowledge of which among correlated variables are cause and which effect.

Summary

1. In partial correlation, the true measure of contribution of a cause to an effect is mutilated if we have rendered constant variables *which may in part or in whole be caused by either of the two factors whose true relationship is to be measured, or by still other unmeasured remote causes which also affect either of the two isolated factors* (*i.e.*, cause and effect).

2. The distinctive contribution of a factor to the magnitude of a multiple correlation coefficient is not a measure of its true contribution to the criterion if any of the other factors entering into the multiple correlation can be described by the clause in italics above.

3. Partial and multiple correlation techniques have been used in studies of causation far more widely than logic would warrant.

References

1. Bagley, W. C.: "Determinism in Education." Baltimore, 1925.
2. Buckingham, B. R.: Partial Correlation. *Journal of Educational Research*, Vol. VII, 1923, pp. 344–349.
3. Burt, Cyril: "Mental and Scholastic Tests." London, 1921.

4. Gates, A. I. and LaSalle, Jessie: The Relative Predictive Value of Certain Intelligence and Educational Tests. *Journal of Educational Psychology*, Vol. XIV, 1923, pp. 577–539.
5. Holzinger, Karl and Freeman, F. N.: The Interpretation of Burt's Regression Equation. *Journal of Educational Psychology*, Vol. XVI, Dec., 1925, pp. 577–582.
6. Kelley, T. L.: Partial and Multiple Correlation. *In* Rietz, H. L. *et al.*, Handbook of Mathematical Statistics, pp. 139–149, 1924.
7. Knight, F. B.: Qualities Related to Success in Teaching. *Teachers College Contributions to Education*, No. 120, N. Y., 1922.
8. McCall, W. A.: "How to Experiment in Education." N. Y., 1923.
9. Reavis, G. H.: Factors Controlling Attendance in Rural Schools. *Teachers College Contributions to Education*, No. 108, N. Y., 1920.
10. Somers, G. T.: Pedagogical Prognosis. Predicting the Success of Prospective Teachers. *Teachers College Contributions to Education*, No. 140, N. Y., 1923.
11. Thomson, G. H.: "Instinct, Intelligence and Character." N. Y., 1924.
12. Whitney, F. L.: The Analysis of Teaching Functions. *Journal of Educational Research*, Vol. VII, 1923, pp. 297–308.
13. Whitney, F. L.: Application of the Analysis and Prognosis of Teaching Functions to Problems of Teacher-training. *Journal of Educational Research*, Vol. VIII, 1923, pp. 20–30.
14. Whitney, F. L.: The Prediction of Teaching Success. *Journal of Educational Research Monograph*, No. 6, pp. 1–85, 1924.
15. Wright, Sewall: Correlation and Causation. *Journal of Agricultural Research*, Vol. XX, 1920–21, pp. 557–585.
16. Yule, Udney: "Introduction to the Theory of Statistics." London, 1924.

(*To Be Concluded in the December Issue*)

THE RELATION OF IQ TO POSITION IN FAMILY[1]

GRACE ARTHUR

Child Guidance Clinic, St. Paul, Minnesota

For three consecutive years it was part of my work in the public schools of Chisholm, Minnesota, to give Kuhlmann-Binet examinations to all the pupils of three kindergartens. In cases where a given family was represented in these kindergarten groups for two of the three years, it thus became possible to compare the intelligence ratings of two members of the same family at approximately the same chronological age and with the same amount of school training. A large percentage of these children were of foreign birth or parentage. The factor of school training is of importance, therefore, as most pupils entered kindergarten with little or no knowledge of the English language. A year in kindergarten, with the special effort that is made by the teachers to encourage the children to use English, is sufficient to enable the majority of them to do justice to themselves when tested with the Kulhmann 1917 Revision of the Binet Scale. This revision appears to be adequately standardized for kindergarten and first grade levels of intelligence, inasmuch as groups with a median mental age of six years in September earn a median score of four at the end of the school year on the Haggerty primary examination in reading comprehension. This is the standard norm for the end of Grade I on that scale. Moreover, groups sorted on the basis of results from the Kuhlmann revision obtained at kindergarten and first grade level have worked together satisfactorily through the completion of Grade VIII.

An inspection of the school system as a whole had given the general impression that the younger members of foreign families were better stuff both mentally and physically than the elder siblings. This impression was strengthened by records of school achievement.

In order to substitute specific data for general impression, intelligence ratings were assembled for 92 pairs of siblings, with Finn, Russian, or South European surnames, all of whom had been tested with the Kuhlmann-Binet scale at the end of one year in kindergarten. The average IQ for the elder siblings is 93.05, with a standard deviation

[1] It is with pleasure that acknowledgment is made for the helpful criticism of data offered by Dr. F. Kuhlmann of the Minnesota State Board of Research and by Professors Karl S. Lashley and Donald G. Paterson of the Psychology Department of the University of Minnesota.

of 11.29. For the younger siblings the average IQ is 99.14 with a standard deviation of 10.42. This difference is large enough to be significant for this group of siblings, as there is not one chance in 100 that the direction of difference would be reversed for other groups similarly selected.

$$\left(\frac{D}{PE_d} = 5.6.\right)$$

This difference might be attributed to any one of a number of factors: Better command of English upon kindergarten entrance, owing to the teaching of the younger child by the elder; better attitude toward the test situation on the part of the younger due to the previous experience of the elder; definite coaching of the younger child by the elder, or a true difference in the abilities of the two siblings at the time of testing. If the difference between the two averages is due to the first named factor, greater language handicap on the part of the elder, it would seem that it should tend to disappear with additional school training of the elder sibling.

There are available Kuhlmann-Binet scores for 271 pairs of siblings of kindergarten and grade school age. These include the 92 pairs described above, of which both siblings were tested at kindergarten level, but also include 179 pairs the elder of which had the advantage over the younger of one year or more than one year of additional training in English. The average IQ for the elder siblings is 89.3; that for the younger siblings is 96.9. That is, instead of eliminating the difference, the inclusion of cases with a greater amount of school training tends to emphasize the contrast. If we eliminate from this group all cases bearing an "American" surname, and retain, as in the group tested at kindergarten level, only those cases bearing Finnish, Russian, or South European surnames, we have 245 pairs of siblings. The average IQ of the elder siblings is now 88.27 while that of the younger is 96.26. The advantage of the younger over the elder would appear to increase slightly with chronological age, just as we should expect it to do if the difference in rating at kindergarten level were based upon a true difference in abilities.

However, the matter of selection of cases may affect the data from the elder groups. The 26 pairs of siblings of American surname that were eliminated from the group show an average IQ of 99.4 for the elder siblings, and one of 103.3 for the younger siblings. From these were then eliminated all cases that showed a difference of more than two years between the chronological age of the elder sibling and that

of the younger at the dates of their respective tests. There remained a group of 19 pairs of siblings. For these, the average IQ for the elder is 104, SD 11.31, while the average IQ of the younger is also 104, SD 10.61. The number of cases is here too small for the results to be anything but suggestive, but the figures seem to indicate that for the larger group in which all ages are included, there may enter the factor of selection of cases, but that a group which includes only cases that show a difference of not more than two years in chronological age between siblings at the time of testing, may probably be regarded as representing a fair sampling.

From the 245 pairs of siblings of foreign surname all cases were next eliminated, therefore, in which the difference in chronological age between siblings at the time of testing was more than two years. This left a group of 190 pairs of siblings. For these the average IQ of the elder was 90.1, while that for the younger was 96.9. This difference of 6.8 points in IQ agrees closely with the difference of 6.9 points that was found between the average IQ's of elder and younger siblings of the 92 pairs tested at kindergarten levels.

Apparently, then, the addition of cases in which the elder sibling has the advantage in amount of school training and consequent exposure to an all-English environment does not affect the difference in ability in favor of the younger sibling as denoted by a comparison of IQ's.

In a number of cases, IQ's were felt to be of questionable validity because of language handicap, and re-tests were given. This works to the advantage of the elder siblings, as it was possible to re-test more of them than of the younger siblings, and the corrected IQ is of course made use of in the present report. In the few cases in which both siblings were re-tested the greater gain in IQ made by the younger sibling suggests another angle from which our problem might be investigated.

To eliminate the second and third possible factors in the better achievement of the younger child, those of better attitude toward the test situation and definite coaching of the younger child by the elder, data were sorted out for 36 cases in which the younger sibling was tested before the elder. In the school that supplied these cases, the kindergarten was the first group tested. Consequently, the strangeness of the test situation was felt most strongly by the kindergarten group. For the 36 pairs of siblings for which data were available, the elder siblings have an average IQ of 84.08. That of the younger sib-

lings is 92.07. After eliminating all cases in which the difference in the chronological age of the two siblings at the time of testing is more than two years, there remain 19 pairs of siblings. For these the average of the elder is 86.5 and that of the younger is 94.7. From the relatively low IQ's of both sibs, one is led to infer that some process of selection is operating in the choice of some of these cases that, because of the small number in the group, tends to color results. This selective process probably would be the referral of some of the elder siblings as problem cases, in place of the wholly chance selection that results from testing only all the members of specific classes. But even though the elder sibling should be referred for testing because of some school difficulty, there seems to be no clear reason why the sibling in kindergarten or first grade should not be just as dull or more so instead of brighter, as he seems to be. As the difference in IQ remained the same for unselected groups whether both siblings were tested at kindergarten level, or one was tested when from one to two years older, the possibility of unequal difficulty of the different levels of the scale operating as a factor seems to be largely eliminated.

The question of the teaching of the younger sibling by the elder was treated with scepticism by a six-year-old who was out-distancing a dull elder brother. "For," quoth he, "how *can* Joey teach me what he does'nt know hisself?" It seemed to be a reasonable query.

The phenomenon of decreasing IQ with priority of birth in the families of foreign surname is still further emphasized by data from families of which three or four children were given Kuhlmann-Binet examinations. For 85 families, three children were tested. The average IQ of the eldest is 82.7, that of the next eldest is 90.34, while that of the youngest is 95.34. For 19 families four children were tested. Of these the eldest have an average IQ of 77.2, the next an average IQ of 80.9, the third eldest, an average IQ of 87.4, and the youngest average 93.5. The families for which tests of three and four children are available are selected largely upon the basis of school difficulty of one sort or another on the part of the elder children. This helps to account for the low average IQ of the children of each position within the family, but only tends to further demonstrate the problem of increasing IQ for each group of cases as we progress from eldest to youngest.

To summarize, then, we seem to have a situation in which the younger of two siblings born of immigrant parents has a distinct advantage, from the standpoint of IQ, over the elder sibling. This advantage

cannot be accounted for by any hypothetical inequality of different levels of the scale by which measurements were made, as the difference in average IQ is significant even when both siblings are tested at approximately the same chronological age, that is, at the end of the first year in kindergarten. Neither can it be accounted for simply by a greater knowledge of English on the part of the younger child, as results remain practically unchanged when the elder child, after two or three years of instruction in English is compared with the younger child at the end of one.

The 19 pairs of siblings of "American" surname that had a difference of not more than two years of chronological age between sibs at the time of testing, showed no difference between the average IQ of the elder and that of the younger. However, the number of cases was too small to permit this result to be regarded as anything more than a suggestion of what might be demonstrated by further investigation. In order to determine what the relationship is between the IQ's of elder and younger siblings of "American" birth and parentage, an effort was made to obtain data for a larger number of unselected cases.

Through the courtesy of Miss Edith Taylor, Principal, the test files of the Hancock School, St. Paul, were put at my disposal. With the help of Miss Nina Whitman who gives a Kuhlmann-Binet examination to all first grade entrants of that school, siblings were grouped together. We found 70 pairs of siblings of whom each had been given a Kuhlmann-Binet test some time between the chronological ages of 5.0 and 7.0. The average IQ for the elder siblings is 105.9, SD 11.23, while the average IQ for the younger siblings is 106.9, SD 10.67. This difference of one point is not highly significant, as there are only 71 chances out of 100 that the difference would occur again in the same direction with a similarly selected group.

From these results it appears that the younger child of an immigrant family has an advantage from the standpoint of IQ over the elder sibling that is not equally shared by the younger of two American siblings. In an effort to discover some indication of the basis of this superiority it seemed desirable to obtain data as to physical development. Mr. J. P. Vaughan, Superintendent of Schools of Chisholm, and Dr. Graham, School Physician, kindly consented to the use for this purpose of the school records of physical examinations. The record of the first physical examination made by the doctor after kindergarten entrance supplied the data used for each case. Direct comparison of height and weight of the two siblings was not considered

advisable, owing to the variation in the date of measuring different groups. It was thought, however, that weight-height coefficients might be significant. When these were calculated and averaged for the 67 pairs of siblings for whom data were still available at the time of my request, the difference between the average for the elder and that for the younger sibling proved to be wholly insignificant. There had also been recorded by the physician an estimate of the general state of nutrition of the child at the time of the physical examination. These grade from good to very poor. Here, the younger sibling is found to have an average rating somewhat higher than the elder, but not significantly so.

So far as these measurements are concerned, then, the advantage of the younger child cannot be explained in physical terms. It would be interesting to see what more extensive measurements, made with this specific question in mind, might reveal.

Acquaintance with certain individual cases may indicate in what direction it may be worth while to look for light. Certain immigrant families fail to show any tendency to vary in IQ according to position in family. These appear to be largely of two types: First, those who have adopted American customs and habits of living in a high degree, that is, whose habits of living seem to have become stabilized at the American level; and second, those whose habits of living have been influenced, relatively, very little by the new environment, that is, who are still stabilized at the earlier level. The mothers of the latter still consider black coffee and home made wine on the list of approved baby foods, and it is not unusual for the teacher of a child from such a home to be obliged to put her small pupil through a "sobering up" process before the work of teaching can begin. These statements are supported by so small a number of cases that they can be regarded as nothing more than suggestions. The cases have illustrative value in that they show what may be happening, but there are not enough of them to demonstrate convincingly what is happening. If a larger number of cases should prove to support this suggestion, however, we would have a situation in which families stabilized at the American level of living exhibit the same tendency toward stability of IQ as American families, and families stabilized at their old world level of living show equal stability of IQ, while families going through a transition stage from their old world habits of living to new world standards tend towards an increasing IQ for successive offspring that works to the advantage of the younger children.

It is perhaps worth while to note that after the age of kindergarten entrance and the reduction of language handicap, the IQ's of these children appear to be as stable as those of any other group with which I have worked. A correlation of 275 IQ's obtained with the Kuhlmann-Binet scale over a period of three years with IQ's for the same individuals obtained with the Arthur and Woodrow revised point scale given near the beginning of the three year period yielded a coefficient of $0.92 \pm .006$ (linear regression). So high a correlation would hardly be possible with all the chance factors that operate in connection with the giving of a group test, unless IQ's were as stable as those of other groups that have been used in reporting group test results. Moreover, in re-testing cases to verify IQ's where their validity was in question because of either supposed language handicap or emotional factors at work during the original examination, the change in IQ is large enough to constitute a significant difference in only a surprisingly small number of cases.

The subject of variation in IQ according to the position in the family group has been treated in some detail by a number of investigators. Dr. Kuhlmann has compiled the accompanying table from the data of Piper, Koenig, and Kerlin, to show the incidence and percentage of feeblemindedness for each position in the family group. He has kindly permitted me to copy it from a manuscript that is as yet unpublished. It is evident that for the groups cited, the highest incidence of feeblemindedness occurs among the first-born and tends to decrease for the succeeding positions. This would of course be expected, other things being equal, as there must of necessity be more children in the position of first-born than there are of second, third, or fourth, so long as the "only child" exists. Turning to the line of

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
Number for each order	338	334	315	292	245	203	155	125	88	56	43	28	20	18	13	8	5	2	2	Piper
	185	183	174	150	127	105	80	60	52	32	20	12	5	3	3	2	2	2	1	Koenig
	100	95	88	75	55	48	40	26	18	10	6	4	1	...	....	....	..	..	..	Kerlin
Total.....	623	612	577	517	427	356	275	211	158	98	69	44	26	21	16	10	7	4	3	
Number mental defectives for each	86	57	73	43	23	27	17	19	4	6	1	3	..	1	2	1	..	..	..	Piper
	39	40	28	23	14	10	8	8	9	5	...	..	..	...	1	....	..	..	..	Koenig
	36	22	21	8	13	8	5	5	3	....	1	...	..	...	....	....	..	..	..	Kerlin
Total.....	161	119	122	74	50	45	30	32	16	11	2	3	..	1	3	1				
Per cent defective...	25.8	19.4	21.1	14.3	11.7	12.6	10.9	15.2	10.1	12.2	2.9	6.8	..	4.8	18.8	10.0				

percentage, however, we note that a higher percentage, also, of these first-born children are feebleminded than of children in any other position in the family, and that there is a definite though somewhat irregular tendency for this percentage to decrease from one position to the next. From the standpoint of incidence of feeblemindedness, then, for groups that have not undergone any radical change in environment, the position of first-born is the least desirable, while succeeding positions tend to become more so.

On the other hand, Terman in his *Genetic Studies of Genius* finds "the data on order of birth, as far as they may be considered valid, are in striking agreement with Cattell's figures in showing a preponderance of first-born gifted in families of two or more. The fact that superiority of the first-born registers in childhood as clearly as in the achievements of adult life suggests that the causes are to be sought in native endowment rather than in environment and education." It is possible that the conclusions of Cattell, Kuhlmann and Terman are all supported by the fact that the standard deviations are consistently greater for the average IQ's of the elder siblings of every group considered in this study, in spite of the lower average IQ's.

The report of Boas as to changes in the character of children of immigrants is all in terms of anthropological measurements. From this standpoint he seems to be convinced that something radical happens to the physical make-up of the children as a result of the radical change of environment involved in moving the immigrant from his old world home to the new. Data from tables of this report are often hard to evaluate because of lack of information as to number of cases, standard deviations, etc. for various groups. But the conclusions based upon so wide a series of observations seem worth considering.

An interesting contribution to the subject of the effect of environment during the pre-school period consists of the published reports of some of the various nursery schools and child study stations that have come into existence during the past decade. Johnson reports temporary spurts in IQ following intensive training in speech and motor control that tend to disappear with later re-tests. The Fifth Annual Report of the Merrill-Palmer School indicates the same tendency for IQ to increase under special training, but no relapse. Baldwin and Stecher seem to find permanent increase in IQ over the period of observation, but they are quite conservative in their interpretation of the phenomenon.

SUMMARY

For 92 pairs of siblings of Finn, Russian and South European surname, all tested with the Kuhlmann-Binet scale, the difference between the average IQ of the elder siblings and that of the younger is 6.8 points. This difference is significant, as there is not one chance in 100 that the direction of difference would be reversed for other groups similarly selected.

The hypothetical factor of unequal difficulty of different parts of the scale was eliminated by testing both siblings at the end of their first year in kindergarten.

The difference found cannot be attributed wholly to greater language handicap on the part of the elder sibling, as it tends to remain constant when 151 additional pairs of siblings are included, the elder of which had the advantage over the younger of one or two years of additional training in English.

The difference cannot be due wholly to the fact of the elder sibling being tested first, as the difference is not reduced by testing the younger sibling first.

For 70 pairs of siblings of "American" surname of whom each had been given a Kuhlmann-Binet examination sometime between the chronological ages of 5.0 and 7.0, the difference between the average IQ of the elder and that of the younger is 1.0. This difference is not highly significant, as there are only 71 chances out of 100 that the difference would occur again in the same direction with a similarly selected group.

It is worth noting that for all groups considered, of both American and foreign surname, the standard deviation of the IQ's of the elder siblings is greater than that of the younger. This tends to support the conclusion of Cattell, Kuhlmann, Terman and other investigators as to the greater incidence of both superiority and inferiority among first-born.

The one outstanding fact, from the practical standpoint, appears to be that, for the community investigated, the elder siblings of immigrant families are, on the average, inferior to the American population in general intelligence, but the younger siblings are, on the average, good intellectual stuff.

The questions that present themselves for further research are: (1) Will these differences continue into adulthood? Our data seem to indicate that they tend to increase rather than to diminish with chronological age.

2. Is their basis to be found in physiological changes in the parents, resulting from a radical change of environment, as indicated in the work of Boas?

3. Or is their basis to be found in the physiological effects of the environment upon the children during the period of gross brain development?

4. Or is it, like the increase in IQ of the children of the child study stations, to be found in a more stimulating intellectual environment?

5. How can we investigate these separate factors?

A NOTE ON SCALING TESTS

GODFREY H. THOMSON

Moray House, University of Edinburgh

In this *Journal* for October, 1925, Thurstone discusses a method of scaling psychological and educational tests, and in the course of his paper he proceeds to ascertain the chronological age at which a test item is at *par*, that is the age at which 50 per cent of children can pass the test question. He illustrates on Burt's London Binet data, and gives (on his page 444) curves of these data. He points out that the most refined method of obtaining the 50 per cent point, by fitting equations, is too laborious, and then he adopts a plan which depends only upon the two percentages which bracket the 50 per cent point, interpolating between them.

This plan discards all the other points of the curve: and the object of the present note is to call attention to an easy method of finding a value which uses *all* the data. The problem is one which is very familiar in psychophysics.[1]

Take in illustration Burt's data for the missing features test:

Age	Per Cent Passing
4½	9.5
5½	34.0
6½	65.3
7½	87.3
8½	95.6
9½	98.6

The simple method proposed by Thurstone (which would of course often be quite sufficient) is to interpolate between 34.0 per cent (at 5½ years) and 65.3 per cent (at 6½ years).

$$p_1 = 34.0$$
$$p_2 = 65.3$$

$$\text{Age at par} = 5\tfrac{1}{2} \text{ years} + \frac{50 - 34.0}{65.3 - 34.0}$$
$$= 5\tfrac{1}{2} + 0.51 = 6.01 \text{ years.}$$

The plan here proposed, for use if it is not desired to discard all the data except these two points, is as follows. It can only be used when

[1] See "The Essentials of Mental Measurement." Brown and Thomson, Cambridge, England, 1925, Chapter III.

it may be presumed, or is known, that practically no children can answer the item at the next age below (here 3½ years) and practically all at the next age above (here 10½ years), and when the points are equidistant (as here a year apart). *Simply add all the percentages together:*

Age	Percentage
4½	9.5
5½	34.0
6½	65.3
7½	87.3
8½	95.6
9½	98.6
	390.3

The point required is then

$$\frac{390.3}{100} \text{ years}$$

below a point half a year (or half a unit) beyond the top age, *i.e.*, below 10 years:

$$9\tfrac{1}{2} + \tfrac{1}{2} - \frac{390.3}{100} = 6.097 \text{ years.}$$

This value it will be seen depends on all the data, and is hardly more laborious, if at all, than interpolation.

Strictly speaking it is not the 50 per cent point but the average of the ages at which the children pass the item. If we take the differences of the percentages, and center them at the mid-points between 4½, 5½, etc., we get the same result, thus:

9.5	− 0.0	= 9.5	at 4	giving	38.0
34.0	− 9.5	= 24.5	at 5	giving	122.5
65.3	− 34.0	= 31.3	at 6	giving	187.8
87.3	− 65.3	= 22.0	at 7	giving	154.0
95.6	− 87.3	= 8.3	at 8	giving	66.4
98.6	− 95.6	= 3.0	at 9	giving	27.0
100	− 98.6	= 1.4	at 10	giving	14.0
	100.0				609.7

$$\frac{609.7}{100} = 6.097 \text{ years.}$$

This last is only given in explanation, not as the method recommended for calculation. The best method for calculation is, as set out earlier, to add together the percentages, divided by 100, and subtract from half a year beyond the top age at which some children still fail.

But although this is not strictly the 50 per cent point, it is the same thing if the histogram just set out is symmetrical, *i.e.*, the histogram 9.5, 24.5, 31.3, 22.0, 8.3, 3.0, 1.4. It is very nearly so, though not quite. Most, if not all, practicable equation-fitting methods either assume a symmetrical curve, or make other assumptions which are less probable.[1]

A slight extension enables *sigma* to be calculated for the item, thus (repeating the first calculation for clearness, and continuing):

Age	Percentage	Continued Sum
4½	9.5	9.5
5½	34.0	43.5, *i.e.*, 9.5 + 34.0
6½	65.3	108.8, *i.e.*, 43.5 + 65.3
7½	87.3	196.1
8½	95.6	291.7
9½	98.6	390.3 (checks with first total)
	390.3	1039.9

$d = 3.903$ $S = 10.399$

Par = top age $9\frac{1}{2} + \frac{1}{2} - d = 6.097$ years.

$\sigma^2 = 2S - d\,(1 + d) - 0.083$ (Sheppard's correction)

whence $\sigma = 1.26$ years.

[1] Fitting a normal integral (the $\phi(\gamma)$ process) gives 6.0697 years.

THE TECHNICAL VOCABULARY OF PSYCHOLOGY

CHAS. L. HARLAN

Lewistown, Idaho

The field of psychology contains a relatively large number of technical terms. Moreover there are many ordinary terms which are used with technical meanings in this subject. A check for several years on the number of words recognized by psychology students as new to them, or as having new meanings for them, reveals approximately 200 technical terms which must be "learned" by students of psychology before they have finished their study of this subject. Such words as neurone, stimulus, reaction and synapse, are used with such frequency and are representative of concepts of such importance that they are fundamental to an understanding of the subject.

One of the chief difficulties encountered by the student beginning the study of psychology is the vocabulary difficulty. New words or words which have no very clear connotation or words used in a technical sense are, at first, a serious handicap to the student in his reading of psychology, or in any psychological discussion. To many beginners in the field the large number of technical terms encountered constitutes, for them, almost a foreign language. If these are not mastered as the student proceeds in his study of this subject, he is soon lost in a seemingly hopeless jargon. He labors over the printed page "sliding over" the technical words which are often the key to the sentence or paragraph, thereby missing some of the fundamental concepts and ideas. Or, he listens to the discussion of his psychology instructor speaking in psychological terms and fails to comprehend "what it is all about." Some careful attention to vocabulary building seems, therefore, to be essential to the effective teaching of psychology.

The writer, with the help of his students in psychology, during a period of six years, has attempted to select: first, those words which are essential to an understanding of psychology; second, those words which cause difficulty for the beginning student in the field of psychology. The method employed was to have each student make a glossary of technical terms during his two quarters of study of this subject. The sources included five widely used beginning psychology texts, eight texts used as reference and any words not already listed which were used in class discussions. In all, more than 400 words were listed, but many of these had so low frequency of mention that

they were considered unimportant, either to an understanding of psychology, or that they presented little difficulty to the beginning student. There were 176 words listed, however, with sufficient frequency to justify their inclusion in the psychology vocabulary of the well informed student.

The next step in the study was to determine in some objective and reliable way how many of these words the student has made a part of his psychology vocabulary. This was done by putting each word in an exercise involving judgment as to the best word in a series of five to complete a statement of meaning, the correct one to be underlined. For example:

The earliest period in human life is:

maturity, adolescence, senility, infancy, youth.

No attempt was made to formulate elaborate technical definitions. The exercises constitute merely statements of the essential meaning of the terms.

The preliminary form of the tests consisted of 160 of these exercises. These were given to three classes enrolling, in all, 90 students who had never had any psychology, and again after they had had two quarters (18 weeks, five hours per week) of psychology. In the results, after two quarters of psychology, were included the scores of 80 students who did not take the tests before studying the subject. On the basis of these results the difficulty of the exercises was determined and the test exercises regrouped into three tests of approximately equal difficulty, each test consisting of 50 exercises. In this form the three tests were then given to approximately 500 students at the end of their semester's work in psychology in seven different normal schools. On the basis of these results tentative norms have been secured and the validity and reliability of the tests determined. Test III has not yet gone through the final stages of standardization, hence the results presented herewith are for Tests I and II only.

As above stated, the 100 exercises from the original list were given to a group of students before they had had any psychology and again to the same students after they had taken two quarters, or 18 weeks of psychology. The results are most interesting. The "before psychology" median was 51 points and the "after psychology" median was 85 points, out of a possible 100, the median gain being 32 points. The largest gains were made by those who made the low scores on the test before having psychology. The correlation coefficient (Pearson method) for gain, in points with total scores on the test before psy-

chology, is −.69. This is an unusually high negative correlation. It means that those who had the smallest psychology vocabulary at the beginning made the greatest gains during their 18 weeks study of psychology, and that those who already had a large vocabulary at first, made the smallest gain. In no case were perfect scores made on the test "after psychology," hence the "limit of possibilities for gains" is not the only explanation of the high negative correlation. Although the correlation of "before scores" with gains on second tests is negative, the correlation of "before psychology" *scores* with "after psychology" scores is high and positive. ($r = +.72$.)

Test I shows a median of 41 (499 cases) as compared with a median of 43 (432 cases) for Test II. The medians for Test I, in the eight different normal schools, range from 38.9 to 44.2, while those of Test II range from 41.7 to 46. The correlation of Test I with Test II (430 cases) is +.76.

The median total score for Test I combined with Test II is 85 and ranges from 79.4 to 89.1 in the different schools. For those who have had 18 weeks or less of psychology (352 cases) the median is 82.5 and for those who have had 24 weeks or more (147 cases) the median is 88.3, a difference of nearly six points in favor of those who have studied the subject a longer time. This is not greatly significant, however, as the gain during the first 18 weeks is 32 points. This would indicate that the process of vocabulary building in the study of psychology takes place very largely in the early stages of the study of the subject.

Scores on Test I correlate with semester grades in psychology +.54. The correlation between Test II scores and semester grades in psychology is +.56 (286 cases). The correlation of total scores—Test I and Test II combined—with semester grades in psychology is +.61. These indicate that Test I and Test II are about equally reliable, and that comprehension of technical terms in psychology is not the only factor in determining semester grades, although that is somewhat important as shown by the high correlation (+.61). In some schools the correlation between semester grades and total score ran much higher than in others, the highest coefficient being +.71.

There was sufficient range in ages of the 500 students tested to arouse some curiosity as to the closeness of the relationship between age and total vocabulary score. For 178 cases the correlation was −.17. This is significant in that it is negative, although not high. There is a tendency, in these tests, for the younger students to make the higher scores.

Intelligence is always a factor in determining performance in any test. It plays its part in influencing the results of the psychology vocabulary tests. The correlation between total scores in this test and total score on the Terman Group Test of Mental Ability, Form A (100 students) is +.71. This correlation is based on the vocabulary test results after the students had had 18 weeks of psychology. The correlation of intelligence with vocabulary before they had had any psychology is only .63. These coefficients are both sufficiently high to indicate the importance of the relationship between intelligence and vocabulary attainment.

It is obvious at the very first consideration that a vocabulary of technical terms is not all of psychology. There is a large body of information both general and specific, together with attitudes of mind and methods of procedure, which every student should possess. There are possibly other outcomes to be secured by the students of this subject. Vocabulary is, however, one of the first difficulties encountered by beginning students of the subject and should, therefore, be taken into account by the instructors in this field. The psychology vocabulary test described above should serve a useful purpose in checking upon the results in this particular bit of the field.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY C. O. MATHEWS
Research Associate, Institute of Educational Research, Teachers College
Columbia University

Educational Psychology

A Century of Applications of Psychology to Education. Charles H. Judd. Teachers College Record, May, 1926, 771–781. A review of the past century of educational psychology and an evaluation of present trends in the field.

Educational Psychology. V. A. C. Henmon and J. A. Melrose. The Psychological Bulletin, July, 1926, 353–365. Bibliography and brief summary of the literature in the field of educational psychology for the last three years.

Mental Testing

Intelligence Tests. Rudolph Pintner. The Psychological Bulletin, July, 1926, 366–381. Bibliography and summary of the investigations dealing with intelligence tests published during the last few years.

The Use of Intelligence Tests by Universities. Herbert A. Toops, School and Society, July 17, 1926, 87–88. The information obtained by a questionnaire to 110 colleges and universities is summarized.

A Group Point Scale for the Measurement of Intelligence. Grace Arthur. The Journal of Applied Psychology, June, 1926, 228–244. The distinctive feature of the scale is the use of a formula to determine the discriminative value of each test used, to reduce scores to comparable units, and to furnish a system of weighting independent of subjective evaluation.

The Mental Age Concept. L. L. Thurstone. Psychological Review, July, 1926, 268–278. This is a criticism of the use of the mental age concept and a suggestion for the substitution of either percentile or sigma standing in its place.

A Study of Some Commonly Used Performance Tests. Myrtle Raymaker Worthington. The Journal of Applied Psychology, June, 1925, 216–227. The correlations between 11 commonly used performance tests and the Stanford-Binet are reported.

An Analysis of Scores on the Illinois Examination. L. H. Lasnier. Peabody Journal of Education, May, 1926, 332–339. The study points out the extent to which abilities (arithmetic and reading) tested by different tests are identical.

The Factor of Speed in Intelligence. Helen Peak and Edwin G. Boring. Journal of Experimental Psychology, April, 1926, 71–94. The conclusion is reached as a result of a laboratory experiment that "speed of reaction is an important, and probably the most important, factor in individual differences in the intelligent act."

The Accuracy of Group Test Mental Ages and Intelligence Quotients of Junior High School Pupils. Fowler D. Brooks. The School Review, May, 1926, 333–342. The Stanford-Binet and nine group tests were given to 108 junior high school pupils. The accuracy of their ratings are shown and their uses in classification and sectioning are discussed.

A New Approach to the Measurement of the Intelligence of Young Children. Florence L. Goodenough. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1926, 185–211. A study of the intellectual factors involved in the spontaneous drawings of children and a scale to measure these factors.

On the Variability of Boys and Girls from Thirteen to Eighteen. Staff of the Division of Psychology of the Institute of Educational Research, Teachers College, Columbia University. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1926, 182–184. Measures of variability of the intelligence of 5300 boys and 6300 girls 13 to 16 years of age show boys to be slightly more variable than girls.

Sex Differences in Status and Gain in Intelligence Scores from Thirteen to Eighteen. Staff of the Division of Psychology of the Institute of Educational Research, Teachers College, Columbia University. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1926, 167–181. Data of the study show that the doctrine of adult male superiority, delayed in its attainment, is quite unlikely. For all practical school purposes sex differences in intelligence may be ignored.

A Study of Natio-racial Mental Differences. Genetic Psychology Monographs, May and July, 1926, Vol. 1, Nos. 3 and 4. A study of national and racial differences among children of school age. Results from intelligence tests, achievement tests and anthropomorphic data and information bearing upon relation of vocation are included.

A Diagnostic Study of the Mathematical Abilities of a Selected Group of College Entrants. Frank C. Touton and Esther Jeffery Terry. California Quarterly of Secondary Education, April, 1926, 275–306. This is a study of the results of 504 college students with the mathematical operations involved in the Thorndike Intelligence Examination for High School Graduates.

The Psychological Test Program. L. L. Thurstone. The Educational Record, April, 1926, 114–126. Norms are given for the psychological examination used in the comparative study by the American Council on Education. The average score is shown for each college on each test.

Achievement Testing

Student Opinion of Type of Examination. Nira M. Klise. School and Society, July 3, 1926, 23–24. Preferences of 337 students are summarized and the principal reasons for favoring the new type of examinations are stated.

Do New and Old Type Examinations Measure Different Mental Functions? Donald G. Patterson. School and Society, August 21, 1926, 245–248. The old type of examination in an orientation course correlates as closely with the new type as it does with itself.

Educational Tests. V. A. Jones and W. A. McCall. The Psychological Bulletin July, 1926, 382–394. A summary of recent trends in educational testing with a bibliography.

Norms for the Group Four Test. Elliott R. Downing. School Science and Mathematics, June, 1926, 638–643. A description of a test in science with norms based on 3201 students.

Iowa Placement Examinations. George D. Stoddard. School and Society, August 14, 1925, 212–216. A description of the construction and a discussion of the uses of the Iowa Placement Examinations.

Handwriting Survey to Determine Grade Standards. John G. Kirk. Journal of Educational Research, April, 1926, 259–272. The results of a survey covering samples from 2000 employees in 21 large concerns are reported. Occupational and general averages and grade standards on the Ayres scale are given.

Senior High School English as Revealed by a Standard Test. E. J. Ashbaugh. Journal of Educational Research, April, 1926, 249–258. A report of the achievement of 245 girls and 65 boys on Tressler's English Minimal Essentials Test in 17 counties of Central Ohio.

The Unfluence of Certain Social Factors upon Scores in the Stanford Achievement Tests. S. A. Courtis. Journal of Educational Research, May, 1926, 311–324. This is a study dealing with the influence of an environmental factor (Sunday School attendance) upon scores on Stanford tests.

Diagnosing Cases of Failure in Algebra. Peter L. Spencer. The School Review, May, 1926, 372–376. A method for the construction and use of a diagnostic test in algebra is suggested.

The Wisconsin Tests of Grammatical Correctness. S. A. Leonard. The English Journal, June, 1926, 430–442. This is a completion type of test useful in finding out the proportion of pupils making certain types of grammatical errors. Norms are given.

The Wisconsin Tests of Sentence Recognition. S. A. Leonard. The English Journal, May, 1926, 348–357. The construction of the tests, their purpose and the average scores of pupils in grades VII to XIII (first year normal school) are given.

Some Achievements of Pupils of the Same Mental Ages but Different Intelligence Quotients. F. T. Wilson. Journal of Educational Research, June, 1926, 43–53. Comparison of achievements in reading, arithmetic operations and problems, and spelling as measured by the Stanford Achievement Tests.

A Critical Study of the Results of the Barr's Diagnostic Tests in American History. Herman M. Wessel. The Historical Outlook, May, 1926, 230–231. The results of a group of 35 pupils who took the test are utilized to point out the types of errors made and the relationship between abilities involved.

Forecasting from Aptitude Examination Scores. John D. Clark. Journal of Chemical Education, July, 1926, 785–786. The relationship between success in Chemistry and score on the Iowa Placement Examination, series CA-1, is shown.

The Influence of Certain Social Factors upon Scores in the Stanford Achievement Tests. S. A. Courtis, Journal of Educational Research, June, 1926, 33–42. Religion was selected as a social factor whose distinctive effects upon achievement it was proposed to determine.

Psychology of Learning

Comparative Effectiveness of the Lecture-demonstration and Individual Laboratory Method. Fred G. Anibel. Journal of Educational Research, May, 1926, 355–365.

The preponderance of evidence from this study favors the lecture-demonstration method.

Specific Serial Learning: A Study of Remote Forward Association. Hulsey Cason. Journal of Experimental Psychology, August, 1926, 299–324. Experiments with difficult prose, easy prose, and poetry tend to show that remote forward associations are not formed and learning of items in a different order is not facilitated.

Specific Serial Learning: A Study of Backward Associations. Hulsey Cason. Journal of Experimental Psychology, June, 1926, 195–227. With non-sense syllables backward associations were formed to some extent even though the series was being learned to all appearances in the forward direction. In learning logic materials, it was found that instead of establishing backward associations, inhibiting associations which interfered with learning the same items in reverse order were formed.

Methods and Theories of Teaching Reading Tested by Studies of Deaf Children. Arthur I. Gates. Journal of Educational Research, June, 1926, 21–32. "A presentation of an outline of the theory underlying the practices of teaching the deaf to read and of the hypothesis which is now being subjected to experimental verification."

Psychology of School Subjects

Sex Differences in Handwriting. S. M. Newhall. The Journal of Applied Psychology, June, 1926, 143–150. "The probability of non-'experts' correctly estimating sex from handwriting is of the order 0.55 or 0.60."

Methods of Discovering and Removing Specific Causes of Failure in Ninth-grade Science. D. H. Loree. The School Review, May, 1926, 377–386. Lack of interest and proper study habits seem to be the chief causes of failure. Methods of diagnosing study habits are given. Several case histories are related.

A Grade-placement Curriculum Investigation. Carleton W. Washburne. Journal of Educational Research, April, 1926, 284–292. A summary of the procedure used in classifying the books in *The Winnetka Graded Book List,* taking into account the interest and difficulty of books children read.

Reading Books "Just for Fun." Harvey C. Lehman. The School Review, May, 1926, 357–364. This study indicates the extent to which elementary, high school and university students read when left to their own initiative.

The Construction of a Reading Vocabulary for the Primary Grades. Arthur I. Gates. Teachers College Record, March, 1926, 625–642. The criteria considered in arranging the vocabulary, the method of selecting the words and the uses of the word list are given.

Old and New Methods of Teaching Primary Reading. Edith M. Peyton and James P. Porter. The Journal of Applied Psychology, June, 1926, 264–276. This is an experimental study of two methods of teaching reading, the newer method based on the more recent psychological method proving superior.

Summary of Arithmetic Investigations (1925) *I.* G. T. Buswell. The Elementary School Journal, May, 1926, 692–703. An annoted bibliography of investigations in 1925, with a summary of the investigations dealing with the curriculum, text books and practice exercises in arithmetic.

Summary of Arithmetic Investigations (1925). *II.* G. T. Buswell. The Elementary School Journal, June, 1926, 745–758. Here are summarized investiga-

tions dealing with tests and measurements, nature of arithmetical processes, drill, individual differences, methods of teaching, and history of arithmetic.

An Analysis of the Phonetic Elements in a Basal Reading Vocabulary. Ruth E. Atkins. The Elementary School Journal, April, 1926, 596–606. This is a study of the phonic elements of the 2500 words of most frequent occurrence in Thorndike's *The Teachers Word Book.*

Summary of Reading Investigations (July 1, 1924, to June, 30 1925). III. William S. Gray. The Elementary School Journal, April, 1926. This section of these summaries include investigations concerning the relation of intelligence to achievement, types of class organization, merits of different methods of teaching beginning reading, and reading and vocabulary studies in arithmetic.

Summary of Reading Investigations (July 1, 1924, to June 30, 1925). IV. William S. Gray. The Elementary School Journal, May, 1926, 662–673. Interpretation, extensive or intensive reading, speed of silent reading, children's interests, nature and content, are the subjects covered in this summary.

Teaching Pupils to Read Arithmetic and Other Subject-matter. Olive Gray. The Elementary School Journal, April, 1926, 607–618. A plan for teaching the reading of arithmetic, geography and history is suggested and the success of its use related.

Psychology of Exceptional Children

Application of Two Techniques in Evaluating Some Policies of Dealing with Bright Children. Vernon A. Jones and William A. McCall. Teachers College Record, May, 1926, 825–835. Two studies are described, whose results show that children of superior mental ability advance educationally more rapidly in segregated classes than average children and faster than pupils of same ability when not segregated.

An Information Test Applied to Juvenile Delinquents. Evelyn Eastman. The Journal of Applied Psychology, June, 1926, 202–203. A list of 100 questions in orientation and general information with an analysis of its use on 300 delinquent patients is given.

The Classification of Mental Deficients (Especially of Girls) for a Special Class Center. Nellie E. Powers. The Journal of Applied Psychology, June, 1926, 188–201. This "is a study of the methods and results of the grouping of adolescent girls whose intelligence quotients are between 50 and 90." Replies from 114 cities to a questionnaire concerning the organization and classification of mentally deficient children are summarized.

A Study of the Mechanical Ability of Delinquent Children of the Los Angeles Juvenile Court, 1925. Frances D. Doughtery. The Journal of Delinquency, March, 1926, 293–311. The relationship between mechanical ability and numerous factors such as achievement, intelligence, attitude, method of attack, and physical condition are studied.

Mental Differences among Juvenile Delinquents. Maud A. Merrill. The Journal of Delinquency, March, 1926, 312–323. Mentality of delinquents is studied in its relationship to nationality, type of offence, home influences, and educational status.

Emotional Reactions of Delinquent Boys of Superior Intelligence Compared with Those of College Students. John C. Tjaden. The Journal of Abnormal and Social

Psychology, July-September, 1926, 192–202. The results of delinquents on the Pressey X-O tests are compared with those of 114 college students.

The Grading and Educating of Superior Children. William T. Root. Childhood Education, June, 1926, 455–464. A discussion of the problems involved in educating superior children. Enrichment is advocated as opposed to grade skipping.

Character and Personality

Personality and Character Tests. Mark A. May and Hugh Hartshorne. The Psychological Bulletin, July, 1926, 395–411. Summary and bibliography of articles and books on tests of personality and character. References to ratings are not included.

Slang as an Indication of Character. Gladys C. Schwesinger. The Journal of Applied Psychology, June, 1926, 245–263. Tests of knowledge of the meaning of slang words and phrases were constructed and reproduced. They fail to discriminate between incorrigibles and school boys who were not incorrigible.

The Measurement of Persistence. John J. B. Morgan and Hazel Lucille Hull. The Journal of Applied Psychology, June, 1926, 180–187. A unique wooden maze with grooves is used to measure persistence. An account of its experimental use on 40 cases with a discussion of its validity is given.

A Revision and Some Results with the Moore-Gilliland Aggressiveness Test. A. R. Gilliland. The Journal of Applied Psychology, June, 1926, 143–150. This is a description of the revision of this test and a consideration of its validity.

Character-education and Scientific Method. M. E. Haggerty. Journal of Educational Research, April, 1926, 233–248. This article points out the present status of character study by use of the scientific procedures such as tests, judgments and case studies, as well as its social significance.

The Life Career Motive and Its Effect on High School Work. Grayson N. Kefauver. The School Review, June, 1926, 426–430. An experiment to determine the effect of the life career motive on high school work seems to show that pupils who had made vocational choices did no better school work than those with no vocational choices.

A Study of the Personality of Student Leaders in Colleges in the United States. A. O. Bowden. The Journal of Abnormal and Social Psychology, July-September, 1926, 149–156. Physical, mental and social characteristic of student leaders are indicated. A number of individual cases are cited and types of personality shown graphically.

Measuring Introversion and Extroversion. Edna Heidbreder. The Journal of Abnormal and Social Psychology, July-September, 1926, 120–134. Introversion and Extroversion are measured by the use of Frey's test of 54 traits, both by self ratings and ratings by associates.

Self-ratings as a Means of Determining Trait Relationships and Relative Desirability of Traits. C. S. Yoakum and Grace E. Manson. The Journal of Abnormal and Social Psychology, April, 1926, 52–64. The effects of the time interval, the degree of comprehension of the trait and the desirability of the trait upon self ratings are pointed out.

Characteristics of School Leaders. Otis W. Caldwell and Ruth Wellman. Journal of Educational Research, June, 1926, 1–13. A study of junior high school leaders as to physique, scholarship, intelligence, and personality traits.

MISCELLANEOUS

The Effect of Maturation upon Chicks. Charles Bird. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1926, 212–234. Chicks without practice in pecking first tested when 30 hours old were as accurate, on the average, as a group of equal age which had the practice of 25 test reactions.

Study Habits of High School Pupils. Percival M. Symonds. Teachers College Record, April, 1926, 713–724. This is a report of the study habits of 10 boys of Grade IX after 30 hours of close observation.

Methods of Investigation of Study Habits. Percival Symonds. School and Society, July 31, 1926, 145–152. Six methods are discussed and evaluated in the light of past published studies.

Growth Stages in Play Behavior. Harvey C. Lehman. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1926, 273–288. The rise and training of play activities are very gradual and chronological age divisions are quite arbitrary.

Play Activities of Persons of Different Ages. Harvey C. Lehman. The Pedagogical Seminary and Journal of Genetic Psychology, June, 1926, 250–251. An attempt to discover the activities children engage in when they follow their own inclinations.

An Experiment in Marking. Edward B. Mersereau. The School Review, May, 1926, 348–356. This scheme involves measuring separately in as objective manner as possible four separate abilities.

An Analysis of the Graphic Representation of the Human Figure by School Children. Helen Ann Zesbaugh. The Elementary School Journal, June, 1926, 759–570. The analysis shows the number of boys and girls separately, whose drawing may be grouped according to certain characteristics such as types of figure, showing action, size, proportion, etc.

Comparison of Teacher and Student Estimates of Grades. Paul L. Whitley. School and Society, August 28, 1926, 278–280. Teachers grades are compared with student's self estimates and with student's estimates of their fellows.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY JOHN HOCKETT
The Lincoln School of Teachers College

The "How" of Statistics Made Clear

Statistics in Psychology and Education, by Henry E. Garrett. New York: Longmans, Green and Co, 1926. Pp. XIII + 317. $3.50.

"Lies, damn lies, and statistics." These three stages of progressive meaning are voiced in a clever quip whose point is somewhat justified by the great contemporary increase in erroneous applications of statistical procedure or fallacious interpretations of results. It is obvious that the statistical method, the objective method of interpreting quantitative results, must be accepted, but its employment demands that intelligence and understanding shall replace mechanical and uncomprehending use of these valuable tools. As Professor Woodworth in his introduction to this books says, "The only cure for bad statistics is apparently more and better statistics." *Statistics in Psychology and Education* has been written partially to fulfill this need.

In the inclusion of certain topics the book is not greatly different from many of the other books in the field of elementary statistics which have appeared recently. There is an unusually satisfactory treatment of frequency distributions, measures of central tendencies and their respective uses, measures of variability, graphic methods, the normal curve, and correlation. The topics are presented in a form understandable even by those who have had only a minimum of algebra. The symbolism is consistent and adheres closely to current usage which is gradually crystallizing into a scientific convention. Care has been taken to present clearly formulas and methods of computation; and good problems and exercises, with answers, are given for study and practice. The style is restrained, yet pleasing.

But the author has done much more in his treatment of other topics. His presentation of reliability is unusually complete, but perhaps this very fulness may make it difficult for a beginner. A single treatment of the concept and a list of the specialized formulas might

have improved this section of the book; however, the purpose deserves approval. A good discussion of partial and multiple correlation and regression which are essential in psychological and educational problems is included, making the book a satisfactory one for a really complete first unit course. The final chapter lists some of the more specialized applications of statistical methods to test construction and their results, and is of special value for reference by students at work upon research problems.

Derivations of formulas are not included. Some effort has been made to indicate a few of the limitations of the formulas presented, but a statement of the assumptions underlying their derivations evidently has not been possible. Greater emphasis upon their meanings and limitations would improve a course in statistics, but obviously special treatment of this phase must be left to a second course. There are numerous errata in the book; yet compared to those in the other books in this field they are few, and the author deserves commendation for the thoroughness and care which he must have exercised in its preparation. This praise is the more deserved since it is understood that a second impression will soon appear in which these errors will have been corrected.

The book is a product of the author's immediate teaching experience. It is understandable, reasonably complete, and thoroughly teachable. In a large class of mature students in which the reviewer used it as a text, it met with sincere approbation. In such a work there is promise of decreased abuse of statistical formulas, and increased critical appreciation and interpretation of quantitative results in psychological and educational studies.

Edwin Maurice Bailor.

Dartmouth College.

Professor Morrison Urges Learning for Mastery

The Practice of Teaching in the Secondary School, by Henry C. Morrison. Chicago: The University of Chicago Press, 1926. Pp. VIII + 661. $4.00.

One who has come in contact with Professor Morrison, either directly or through reading his reports and occasional articles, expects whatever he writes to show originality, vigorous thinking, and clarity of expression. His recent book, *The Practice of Teaching in the Secondary School* meets these expectations. It is outstandingly

stimulating alike to the classroom teacher and to the supervising and administrative officer. To the busy teacher the size of the book is appalling; once begun it holds the reader's interest to the end.

The central idea of the book that learning must result in complete mastery has been held by the author for many years. The Laboratory Schools of the University of Chicago, of which he is the director, have given an opportunity to put the idea to a more complete test than he has before found possible. Whatever doubts the administrators of a public school may have as to its applicability to his own situation, he feels as he reads the book that the author is not merely expounding a theory but is describing a practice. The book abounds in concrete suggestions of procedure and citations of cases.

The author attacks with vigor and frequent reiteration the common practice of teaching which results in lesson-learning on the ground-to-be-covered, time-to-be-spent basis, with the accompanying marks presumed to show the degree of the pupil's attainment and concludes that "the normal product of practice in lesson-learning is improvement in the ability to get lessons, and that lesson performance transfers to learning in the real units only casually and in a minority of instances." He would substitute for this procedure what he calls the "mastery formula"; pre-test, teach, test the result, adapt procedure, teach and test again to the point of actual learning." In this plan there is no place for marks of attainment; there are no failures; the mastery of all is complete.

The book is divided into four parts: "Fundamentals in the Teaching Process," "Control Technique," "Operative Technique," "Administrative Technique." Considerably more than half the book is devoted to operative technique. This may be briefly summarized in his "teaching cycle" consisting of five steps, exploration, presentation, assimilation, organization, and recitation. Five types of teaching are recognized, the scientific type, the appreciative type, the practical arts type, the language arts type, and the pure-practice type, and the technique appropriate to each is given in detail. The chapter "Technique in Foreign Language" perhaps reveals the most striking break from the traditional methods of teaching.

It is inevitable that most readers will despair of applying in full the mastery principle with its accompanying technique as set forth in this book. The teachers and supervisors in few schools passes the requisite professional training and insight. A plan that sets aside so completely the established methods of graduation and promotion

requires trained personnel workers for the study of individuals, and teachers skilled in handling corrective and remedial cases not found in the present school staffs.

The author fully recognizes these difficulties. All he hopes to accomplish in this volume, he says, is to give the practicing teacher a method of pedagogical thinking. "No matter how clearly all seem to catch the essential message of the book and no matter how promising the early results may be, the school itself as a whole must grow into it little by little." In the last few pages he has given specific suggestions for the gradual introduction of the plan.

Whether one desires to adopt the mastery principle or not there is much in the book that the progressive teacher or principal can read with profit and find applicable to his work. Parts two and three on control and operative technique abound in such material. Much of this is not to be found in other available literature on method and all is presented with such vigor and from such a novel viewpoint as to make it of unusual interest.

The chapter on "The Integrity of the School" should be read by every principal. With the rapid expansion of the high school, there is danger that the principal will lose sight of guiding aims and will become little more than a performer of routine taks. The following is one of many passages to which the principal should give careful reflection:

"In the hypothetical school consisting of Mark Hopkins and his pupil Garfield, or in the very real Drumtochty School under old Domsie, all these phases of technique coalesce in the personality of the teacher. Unhappily, perhaps, society is long past that possibility. The school is organized and must be organized on the principle of division of labor, but there is all the difference in the world between the school which is a sort of pedantic factory and one which is in some degree an expansion of Mark Hopkins or of the old humanist. In the one, the pupil encounters a series of more or less erudite lesson-hearers, periodically is fed into some testing machine, and occasionally, if he is very good or very bad, has an interview with a person who is classified as an administrative officer. In the other, the whole pupil goes to school to every teacher, to his personnel officer, and to the principal of the school or his assistant. Each of the latter in a school which is at all well organized to secure true educational products is wholly teacher or, if you prefer, educator. True, this person may be using history and another mathematics and a third physics as media of

education, while a fourth is studying the pupil quite apart from any course, and a fifth is occupied with arranging a pupil's schedule or interviewing a teacher or parent. Nevertheless, to the extent that they are all teachers, and teachers in the same sense of the word, the school succeeds or fails in its principal mission."

FRANKLIN W. JOHNSON.

Teachers College, Columbia University.

MORE ABOUT READING

Applied Psychology of Reading, by Fowler D. Brooks. New York: D. Appleton and Company, 1926. Pp. XVII + 278. $1.80.

This book attempts to present the more significant information concerning the psychology of reading in a brief, non-technical form for the teacher. The author seems to be well acquainted with the "vast body of material on the psychology of reading" with which he deals, but in his attempt to be "brief and non-technical" he has done what textbook writers usually do in re-writing material for classroom use—he has run the risk of being uninteresting. The book lacks vitality, punch, force, "kick"—whatever you choose to call it. In my opinion it is encyclopedic, a re-hash of old stuff, a collection of obvious remarks, an admirably dry catalogue of studies and investigations in reading. It is informational—but isn't alive. It lacks the compelling force of actual experience with children, of original experimentation and investigation.

Perhaps the criticism is too severe. The book is no worse than the stuff usually fed to teachers in methods courses in normal schools and teachers colleges. It shows good organization of material, but it is too well grounded in the traditions of the conventional school situation. It reeks of the normal school classroom.

The chapters are preceded by a well-outlined table of Contents and concluded with nine pages of selected references. The chapters on "The Nature and Processes of Reading," "Motor Phases of Reading," "Meaning," "Understanding Sentences and Paragraphs" and "The Measurement of Achievement in Reading" are the best in the book while those chapters dealing with the improvement of reading and remedial measures do not seem very helpful because illustrative descriptions have been sacrificed for terse suggestions and compact principles.

As a mild and limited compendium of references to investigations and studies in reading made up to the present, it will serve a purpose, but as *applied* psychology of reading, as something new or additional in the field of the teaching of reading, it is inadequate.

The last chapter in the book seems to me the worst offender. The whole idea underlying the "appreciation of reading" as something tacked on to "thorough comprehension and rapid rate" is fundamentally wrong. Appreciation is not only "an important aim in reading" —it is the only aim. Comprehension and rate are only tools, another form of manipulation in which the child must acquire skill in order to be able to enjoy that important section of his environment contained in printed pages. "Appreciation" must be thought of as something more than the "attitude of passive enjoyment." It is the re-maker of experience, a prime mover underlying conduct. We may collect and arrange "guiding principles in developing the appreciation of good books" to our academic fill—but meanwhile children await those intrepid pedagogical "Pans" who manage to survive the atmosphere of public school classrooms and teacher-training institutions with enough sympathetic insight into childhood to be bold to inquire, "What is a *good* book?" "Why is a good book *good?*" and who will write books, not only *for* children, but *with* them.

ANNA SHUMAKER.

Teachers College, Columbia University.

HELPING THE FRESHMAN FIND HIMSELF

The Orientation of College Freshmen, by Henry J. Doermann. Baltimore: The Williams and Wilkins Company, 1926. Pp. 158. $3.00

In view of the widespread interest which has been manifested in the orientation of college freshmen during the past decade, it is remarkable that no serious investigation of the movement has heretofore been attempted. Dr. Doermann's pioneer book, therefore, serves to lay a basis of facts which show the origins of the movement, the extent of its present development, and its future trend.

After a preliminary section in which are described the conditions giving rise to efforts toward orientation—namely, high rate of elimination, delinquency, ignorance of college life, etc.,—the author gives an account of the orientation activities in the colleges and universities of the United States. An especially valuable feature in this connec-

tion consists of the outline of orientation courses given in Brown University, Columbia College, the University of Maine, the University of Minnesota, and other institutions.

Maintaining that the ends of orientation can be best met by the organization of a Bureau of Guidance or Personnel Service, the author devotes the last section of his work to a plan for organizing such a bureau, with special reference to the service which it should render to freshmen.

The book is liberally supplied with tables, graphs, and bibliographies and should constitute a serviceable handbook for those interested in this phase of college administration.

H. D. Kitson.
Teachers College, Columbia University.

The Mental Ability of Foreign-born Children

Testing Immigrant Children; Psychological Study of Immigrant Children at Ellis Island, by Bertha M. Boody. Baltimore: The Williams and Wilkins Company, 1926. Pp. 163, 44 tables, bibliography.

The term immigrant has so long connoted a typical adult representative of a race or nationality entering another land that the real force and persistence of individual differences seem to have been lost sight of. It of course has been necessary for legislation to deal with the mass in an effort to "protect society from disintegrating forces," and the practical procedure followed has been one of negative selection rather than being in any way distinctly constructive. However, it seems that one might well expect to find in forthcoming literature some evidence of development.

A psychological study of immigrant children at Ellis Island is more suggestive of progressive approaches than actually being one or than even being conclusive in the report of its findings. It emphasizes a consideration of the individual as such rather than as representative of race or nationality, as well as the need of more appropriate tests. The writer states her purpose as an inquiry into the quantity and quality of people coming to this country.

The report as a whole contains a large amount of reference material and comments concerning immigration in general but very little in regard to children. Two chapters review articles by writers not in immigration work such as Laughlin, Jennings, Salmon, Rosanoff Holmes, and Goddard; and other articles and statistics by writers

connected with immigration, including Mullan, Knox, Gwyn, Nute, Foster, Kerr, and de Moor respectively. They furnish support to a more discriminating selection of the individual. In addition to these subjective as well as objective statements the illiteracy test established in 1917 and the so-called 3 per cent law of 1921 are discussed. A chapter describing the procedure at Ellis Island gives a personal note. The report of the experimental study constitutes the last third of the book.

During August and September, 1922, and July and August, 1923, various psychological tests were given to children in groups and separately who voluntarily attended the school at Ellis Island during detention there. About 100 children the first year and 200–300 the second representing about 20 nationalities participated. The Army Beta, Whipple Cancellation, Johns Hopkins Substitution, Number, and Coordination, Memory for Digits, Dots, Dearborn Form Board, Card Sorting, and selections from the Stanford-Binet and Woolley series were experimented with.

The results of the tests given the children from 3 to 7 years of age with two exceptions of 8 and 9 years are presented in tables. The total number of cases is too small to be differential for age, sex, race, or nationality. The actual tests utilized were the drawing of a square and diamond, the divided rectangle, bow knot, and a comparison of forms from the Stanford-Binet; drawing of a cross, circle and star, button holes, pyramid with blocks, and touching fingers from Mrs. Woolley's tests; and a test of discrimination of color, form, and size. "In this brief study of these youngest representatives of race nothing that sets one kind off from another is evident. Even age within the limits here found does not show itself to a great extent. As they come before you one after the other perhaps the only surprise is that they are so much alike. We have seen in the five Binet tests that in general the age norm held. Of 61 individual successes, 12 were under the age set for the test. Of 72 individual failures only 25 were of the age where according to the Stanford-Binet standards they should have passed."

The conclusions are similar in the case of the results for the 8 to 16 year group to which were given the following tests: Discrimination of color, form, and size, Army Picture Completion test 10, Dearborn Form Board, Binet weight, Card Sorting, Johns Hopkins Number test, Army Beta, and Cancellation. The experiment as might be expected demonstrated the advantages of the non-language method.

The chief value in this report is that it stimulates one to consider the possibilities of those immigrants who are permitted to remain, particularly the children. The report however stops where one would wish it to commence. There should be further study of the entrants. Racial and national traits are nice problems for the anthropologist and psychologist to work out together but it is the growing prospective citizen that should claim the attention of the latter especially. Having apparently eliminated some possibilities of liability thought might be given with profit to the general capacities and particular abilities of immigrant children.

GRACE A. TAYLOR.
Psychologist, Department of Child Guidance, Newark Public Schools.

BEGINNING READERS FOR JOSÉ AND ANITA

The Open Door Primer and *The Open Door First Reader*, by Elma A. Neal and Ollie Peary Storm. New York: The Macmillan Company, 1926. Pp. 98 and 116, respectively.

This primer and first reader are planned for the use of Mexican, Spanish, and perhaps Indian, children who come to school with little or no knowledge of the English language. Manuals provide oral pre-primer lessons and oral work to precede each lesson in the readers. The *Primer* presents the usual *I am* and *This is* start and is padded with considerable static material. The *First Reader* obviously "Amercanizes," often at the expense of ease and childlikeness. There is a feeling of forced and painful effort. Frequent pages of simple questions and commands which may be used for thought-getting in silent reading constitute a good feature of both books.

ANNA SHUMAKER.

OTHER PUBLICATIONS RECEIVED

A. PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY

DODD, STUART C.: *A Statement Concerning Research on an Intelligence Scale for International Use.* Princeton, New Jersey, Princeton University store, 1926, pp. 19.

FREEMAN, FRANK N.: *Mental Tests. Their History, Principles and Applications.* Boston, Mass., Houghton Mifflin Company, 1926, pp. 503.

HELSETH, INGA OLGA: *Children's Thinking.* New York City: Bureau of Publications, Teachers College, Columbia University, 1926, pp. 163.

HOLLINGWORTH, LETA S.: *Gifted Children. Their Nature and Nurture.* New York: The Macmillan Company, 1926, pp. 374

STRASHEIM, J. J.: *A New Method of Mental Testing.* Baltimore: Warwick and York, 1926, pp. 158.

TILTON, J. W.: *The Relation between Association and the Higher Mental Processes.* New York City: Bureau of Publications, Teachers College, Columbia University, 1926, pp. 55.

B. PUBLICATIONS IN PSYCHOLOGY

COSTER, GERALDINE: *Psycho-analysis for Normal People.* New York City: Oxford University Press, American Branch, 1926, pp. 232.

HETZER, HILDEGARD: *Der Einflutz der negativen Phase auf sociales Verhalten und literarische Produhtion pubertierender Mädchen;* and VECERKA, LUCIA: *Das soziale Verhalten von Mädchen währand der Reifezeit.* Jena, Gustav Fischer, 1926, pp. 126.

HOLLINGWORTH, H. L.: *The Psychology of Thought.* New York: D. Appleton and Company, 1926, pp. 329.

LIPPS, THEODOR (translated by SANBORN, HERBERT C.): *Psychological Studies.* Baltimore: the Williams and Wilkins Company, 1926, pp. 333.

MILLER, E. MORRIS: *Brain Capacity and Intelligence.* Sydney, Australia, Australasian Association of Psychology and Philosophy, 1926, pp. 79.

ROBINSON, EDWARD S.: *Practical Psychology.* New York: The Macmillan Company, 1926, pp. 479

SEASHORE, CARL E. (Editor, Psychological Monographs): *University of Iowa Studies in Psychology, No. IX.* Princeton, New Jersey: Psychological Review Company, 1926, pp. 264.

C. PUBLICATIONS IN THE GENERAL EDUCATIONAL FIELD

ANTHONY, KATHERINE M.: *A Manual for Observation and Teaching in the Elementary Grades.* Baltimore: Warwick and York, Inc., 1926, pp. 69.

AVERILL, LAWRENCE A.: *Educational Hygiene.* Boston: Houghton Mifflin Company, 1926, pp. 546.

BARR, A. S. and BURTON, WM. H.: *The Supervision of Instruction.* New York: D. Appleton and Company, 1926, pp. 626.

BENNETT, G. VERNON: *The Junior High.* (Revised Edition.) Baltimore: Warwick and York, Inc., 1926, pp. 225.

BOBBITT, FRANKLIN: *Curriculum Investigations.* Supplementary Educational Monographs, No. 31. Chicago: The University of Chicago, June, 1926, pp. 204

BUCKINGHAM, B. R.: *Supply and Demand in Teacher-training.* Columbus, Ohio: The Ohio State University: March, 1926, pp. 182.

BULL, RAYMOND C. and THOMAS, STANLEY: *Freshman Hygiene.* Philadelphia: J. B. Lippincott and Company, 1926, pp. 228.

COOK, WM. A.: *High School Administration.* Baltimore: Warwick and York, Inc., 1926, pp. 378.

DOUGLASS, HARL R.: *Modern Methods in High School Teaching.* Boston: Houghton Mifflin Company, 1926, pp. 544.

FENTON, NORMAN: *Self-direction and Adjustment.* Yonkers-on-Hudson, New York: World Book Company, 1926, pp. 121.

FRANK, J. O.: *How to Teach General Science.* Philadelphia: P. Blakiston's Son and Company, 1926, pp. 240.

HERTZBERG, OSCAR E.: *A Comparative Study of Different Methods Used in Teaching Beginners to Write.* New York City: Bureau of Publications, Teachers College, Columbia University, 1926, pp. 61.

INSKEEP, ANNIE D.: *Teaching Dull and Retarded Children.* New York: The Macmillan Company, 1926, pp. 455.

JACKSON, CLARENCE M. (Compiler): *Research in Progress at the University of Minnesota, July, 1924–July, 1925.* Minneapolis: Published by University of Minnesota, 1926, pp. 306.

JORDAN, A. M.: *Children's Interests in Reading.* Revised Edition. Chapel Hill, N. C.: The University of North Carolina Press, 1926, pp. 103.

KENNON, LAURA H. V.: *Tests of Literary Vocabulary for Teachers of English.* New York City: Bureau of Publications, Teachers College, Columbia University, 1926, pp. 78.

LOOMIS, ARTHUR K.: *The Technique of Estimating School Equipment Costs.* New York City: Bureau of Publications, Teachers College, Columbia University, 1926, pp. 112.

REEDER, WARD: *The Chief State School Official.* Washington, D. C.: Government Printing Office, Bulletin, 1924, No. 5, pp. 67.

STRANG, RUTH: *Subject-matter in Health Education.* New York City: Bureau of Publications, Teachers College, Columbia University, 1926, pp. 108.

TERRY, PAUL W.: *Extra-curricular Activities in the Junior High School.* Baltimore: Warwick and York, Inc., 1926, pp. 122.

WEST, MICHAEL: *Learning to Read a Foreign Language. An Experimental Study.* New York: Longsman, Green and Co., 1926, pp. 56.

WOODBURNE, A. S.: *Human Nature and Education.* New York: Oxford University Press, 1926, pp. 286

D. NEW SCHOOL TEXTBOOKS

HORN, ERNEST and ASHBAUGH, ERNEST: *Lippincott's New Horn-Ashbaugh Speller.* Philadelphia: J. B. Lippincott Company, 1926, pp. 131.

MCMURRY, FRANK M. and BENSON, C. BEVERLEY: *Social Arithmetic.* Book Three. New York: The Macmillan Company, 1926, 338.

PATCH, EDITH M.: *First Lessons in Nature Study.* New York: The Macmillan Company, 1926, pp. 287.

WILLIAMS, GRACE C.: *The Little Bookshelf.* Chicago: Rand McNally and Company, 1926, pp. 174.

E. OTHER PUBLICATIONS

BUCHHOLZ, H. E.: *U. S. A Second Study in Democracy.* Baltimore: Warwick and York, Inc., 1926, pp. 400.

CRILE, GEORGE W.: *A Bipolar Theory of Living Processes.* New York: The Macmillan Company, 1926, pp. 405.

DELEKAT, FRIEDRICH: *Johann Heinrich Pestalozzi.* Leipzig: Quelle und Meyer, 1926, pp. 314.

LAW, FREDERICK H.: *How to Write and How to Deliver an Oration.* New York: G. P. Putnam's Sons, 1926, pp. 162.

TREAT, KATHERINE: *An Industrial Calendar,* New York City: Vocational Adjustment Bureau for Girls, 1926, pp. 54.

WOODBRIDGE, FREDERICK J. E.: *The Realm of Mind.* New York City: Columbia University Press, 1925, pp. 139